

ORIGIN									
Query Match 94.6%; Score 969; DB 6; Length 4487;									
Best Local Similarity 99.0%; Pred. No. 8.9e-227; Mismatches 5; Indels 5; Gaps 4;									
Matches 1017; Conservative 0;									
Qy	1	ACCGCTCGTCCGCGCGCTTGAGGCGCGCGGAGCGCGCAATTCGTGCGCGCGGG	60						
Db	4	ACCGCTCGTCCG-CGCGCTTGAGGCGCGCGGAG-GCGCGCAATTCGTGCGCGCGGG	61						
Qy	61	GGGGGGGCC--TCCCGGCATCTTCGCGGCACCAAGGACTACAGAGAGGAGGCGGTG	118						
Db	62	GGGGGGGCCATCCCGGCACTTCGCGGCACCAAGGACTACAGAGAGGAGGCGGTG	121						
Qy	119	GGATGGCGGTCGCGGCGCCCGCGAGTACAAAGCGGCGACCTGGTTCGCAAGATGA	178						
Db	122	GGATGGCGGTCGCGGCGCCCGCGAGTACAAAGCGGCGACCTGGTTCGCAAGATGA	181						
Qy	179	AGGGCTACCGCACTGGCGGCGCGGATGATGAATCCAGAGGCGCTGTGAAGCCTC	238						
Db	182	AGGGCTACCGCACTGGCGGCGCGGATGATGAATCCAGAGGCGCTGTGAAGCCTC	241						
Qy	239	CAGCAAAAGTATCTATCTCTCTTTTGGACCCATGAATCTGATTTCTAGTCCCA	298						
Db	242	CAGCAAAAGTATCTATCTCTCTTTTGGACCCATGAATCTGATTTCTAGTCCCA	301						
Qy	299	AAGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACGAAAG	358						
Db	302	AAGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACGAAAG	361						
Qy	359	GATTTAACGAAGATTGTGGGAAATAGAAAATAAACCAGGAGTAAAGTTTACTGGCTAC	418						
Db	362	GATTTAACGAAGATTGTGGGAAATAGAAAATAAACCAGGAGTAAAGTTTACTGGCTAC	421						
Qy	419	AGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGGAGAGGTGGAATATCTGCGATG	478						
Db	422	AGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGGAGAGGTGGAATATCTGCGATG	481						
Qy	479	CAAGCAGTGGGAAGAGGTGATAGATAGAGAGTGAAGAGTGGAAAGGCAAAAGAAAG	538						
Db	482	CAAGCAGTGGGAAGAGGTGATAGATAGAGAGTGAAGAGTGGAAAGGCAAAAGAAAG	541						
Qy	539	AAAAAGCAGGCTCAAAACGGAAGATCATATATCTTCAAGAAATCCTCTAAACAGTCCC	598						
Db	542	AAAAAGCAGGCTCAAAACGGAAGATCATATATCTTCAAGAAATCCTCTAAACAGTCCC	601						
Qy	599	GGAATCTCAGAGATGAAGATGACAAAGCTGCAAGAGAGGAAACAAAGAGCT	658						
Db	602	GGAATCTCAGAGATGAAGATGACAAAGCTGCAAGAGAGGAAACAAAGAGCT	661						
Qy	659	CTGAGGCTGAGATGCGGCAACGACACAGAAACCACTTCAGACTTCGAGAAACCA	718						
Db	662	CTGAGGCTGAGATGCGGCAACGACACAGAAACCACTTCAGACTTCGAGAAACCA	721						
Qy	719	GTGAAGGACCTAACTACCAATATGAATGCTGCATATTAAGAGAAACCAAGAGTT	777						
Db	722	GTGAAGGACCTAACTACCAATATGAATGCTGCATATTAAGAGAAACCAAGAGTT	781						
Qy	778	ATATGTTGGTGTCTAAATCTTGGATTTGATGATGAACCAACATAGCTCTTGTGT	837						
Db	782	ATATGTTGGTGTCTAAATCTTGGATTTGATGATGAACCAACATAGCTCTTGTGT	841						
Qy	838	CATTGACAGAACCCAGTTGTATGATATTAATTTAAAGTTTACTGATTTAATTTTGGTTG	897						
Db	842	CATTGACAGAACCCAGTTGTATGATATTAATTTAAAGTTTACTGATTTAATTTTGGTTG	901						
Qy	898	GGGAAAGACATTTTACGCTTTTAAAGTTTACTGATTTAATTTTGGTTG	957						
Db	902	GGGAAAGACATTTTACGCTTTTAAAGTTTACTGATTTAATTTTGGTTG	961						
Qy	958	CATGAAGTTGCCCTTAAACCACTAAGGATTAACAGATTTTGGCAGACTTATACATGTC	1017						
Db	962	CATGAAGTTGCCCTTAAACCACTAAGGATTAACAGATTTTGGCAGACTTATACATGTC	1021						

Qy	1018	TAGGATC	1024						
Db	1022	TAGGATC	1028						
RESULT 9									
AR240312									
LOCUS									
Sequence 68 from patent US 6468758.									
AR240312									
ACCESSION									
AR240312.1									
VERSION									
GI:27285392									
KEYWORDS									
Unknown.									
ORGANISM									
Unclassified									
REFERENCE									
1 (bases 1 to 1664)									
AUTHORS									
Benson, D.R., Lodes, M.J., Mitcham, J.L. and King, G.E.									
TITLE									
Compositions and methods for ovarian cancer therapy and diagnosis									
JOURNAL									
Patent: US 6468758-A 68 22-OCT-2002;									
FEATURES									
Location/Qualifiers									
1..1664									
source									
/organism="unknown"									
BASE COUNT									
566 a 280 c 319 g 499 t									
ORIGIN									
Query Match 80.9%; Score 828.4; DB 6; Length 1664;									
Best Local Similarity 98.7%; Pred. No. 2.4e-192;									
Matches 835; Conservative 0; Mismatches 11; Indels 0; Gaps 0;									
Qy	179	AGGGCTACCCGACCTGGCGCGCGCGGATTTGATGAATCTCCAGAGGCGCTGTGAAGCCTC	238						
Db	17	AGGGCGCGCCAGCTGCGGGTAGCCCGGATTTGATGAATCTCCAGAGGCGCTGTGAAGCCTC	76						
Qy	239	CAGCAAAAGTATCTATCTCTTTTGGCAGGAGTCAAACTGCAATTTCTAGTCCCA	298						
Db	77	CAGCAAAAGTATCTATCTCTTTTGGCAGGAGTCAAACTGCAATTTCTAGTCCCA	136						
Qy	299	AAGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACGAAAG	358						
Db	137	AAGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACGAAAG	196						
Qy	359	GATTTAACGAAGATTGTGGGAAATAGAAAATAAACCAGGAGTAAAGTTTACTGGCTACC	418						
Db	197	GATTTAACGAAGATTGTGGGAAATAGAAAATAAACCAGGAGTAAAGTTTACTGGCTACC	256						
Qy	419	AGGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGGAGAGGTGGAATATCTGAGATG	478						
Db	257	AGGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGGAGAGGTGGAATATCTGAGATG	316						
Qy	479	CAAGCAGTGGGAAGAGTGTATAGTAGAAGAGATGGAAAGGCAAAAGAGAAATG	538						
Db	317	CAAGCAGTGGGAAGAGTGTATAGTAGAAGAGATGGAAAGGCAAAAGAGAAATG	376						
Qy	539	AAAAAGCAGGCTCAAAACGGAAGATCATATATCTTCAAGAAATCTCTTAAACAGTCCC	598						
Db	377	AAAAAGCAGGCTCAAAACGGAAGATCATATATCTTCAAGAAATCTCTTAAACAGTCCC	436						
Qy	599	GGAATCTCAGAGATGAAGATGACAAAGCTGCAAGAGAGGAGGAAACAAAGAGCT	658						
Db	437	GGAATCTCAGAGATGAAGATGACAAAGCTGCAAGAGAGGAGGAAACAAAGAGCT	496						
Qy	659	CTGAGGCTGGAGATCGGCAACGACACAGAAACCACTTCAGACTTCGAGAAACCA	718						
Db	497	CTGAGGCTGGAGATCGGCAACGACACAGAAACCACTTCAGACTTCGAGAAACCA	556						
Qy	719	GTGAAGGACCTAACTACCAATATGAATGCTGCATATTAAGAGAAACCAAGAGGTTA	778						
Db	557	GTGAAGGACCTAACTACCAATATGAATGCTGCATATTAAGAGAAACCAAGAGGTTA	616						
Qy	779	TATGTTGGTGTCTAAATCTTGGATTTGATGATGAACCAACATAGCTCTTGTGTC	838						
Db	617	TATGTTGGTGTCTAAATCTTGGATTTGATGATGAACCAACATAGCTCTTGTGTC	676						

plwty-9/23/20

Db 901 AAAAGTTACTGATTTAAATTCATGTTATTTGGTGCATGAAGTTGCCCTTAACCACTAAG 960
Qy 983 GATTATCAAGATTTTGGCGCAGACTTATACATGCTCTAGGATC 1024
Db 961 GATTATCAAGATTTTGGCGCAGACTTATACATGCTCTAGGATC 1002

RESULT 7
BD131023
LOCUS BD131023 3805 bp DNA linear PAT 18-SEP-2002
DEFINITION Human growth factor homolog.
ACCESSION BD131023
VERSION BD131023.1 GI:23225968
KEYWORDS JP 2002501747-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3805)
Hillman,J.L., Lal,P., Corley,N.C., Baughn,M., Guegler,K.J. and Tang,T.Y.
TITLE Human growth factor homolog
JOURNAL Patent: JP 2002501747-A 2 22-JAN-2002;
COMMENT INCYTE PHARMACEUTICALS INC
OS Homo sapiens (human)
PN JP 2002501747-A/2
PD 22-JAN-2002
PF 12-JAN-1999 JP 2000529435
PR 29-JAN-1998 US 09/015412
PI JENNIFER L HILLMAN, PREETI LAL, NEIL C CORLEY, MARIAH BAUGHN, KARL

PI J GUEGLER,
PI TOM Y TANG
PC C12N15/09,A61K38/00,A61K39/395,A61K45/00,A61P43/00,
PC A61P43/00,
PC C07K14/475,C07K16/22,C12N1/15,C12N1/19,C12N1/21,C12N5/10, PC
C12P21/02,
PC C12Q1/68,G01N33/53,C12N15/00,A61K37/02,C12N5/00 CC a or g or
C or t, unknown, or other
CC 2675869
FH Key Location/Qualifiers
FT unsure 3788.
Location/Qualifiers
1..3805
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 1255 a 654 c 730 g 1165 t 1 others
ORIGIN
Query Match 95.78; Score 980; DB 6; Length 3805;
Best Local Similarity 99.18; Pred. No. 1.7e-229;
Matches 1017; Conservative 0; Mismatches 5; Indels 4; Gaps 3;

Qy 1 ACCGCTCGTCGGCCGCGCTTTAGCGCCGCGGAGCGCGCAATTCGTCGCGCCGCGG 60
Db 4 ACCGCTCGTCGG-CCGCGTTGAGCCGCGCGGAG-CCGCGCAATTCGTCGCGCCGCGG 61
Qy 61 GGGCGCGCC--TCCCGGCATCTTCGCGGCACCAAGAGACTACCAGGAAGGGAGCGGCTG 118
Db 62 GGGCGCGCCATCCCGCAGCTCTTCGCGGCACCAAGAGACTACCAGGAAGGGAGCGGCTG 121
Qy 119 GGATGGCGCTCCGCGCCCGCGAGTACAAAGCGCGGAGCTGCTTCGCCAAGATGA 178
Db 122 GGATGGCGCTCCGCGCCCGCGAGTACAAAGCGCGGAGCTGCTTCGCCAAGATGA 181
Qy 179 AGGGCTACCGGCACTGGCCCGCGGATTCATGAATCCAGAGGCGCTGTGAAGCCTC 238
Db 182 AGGGCTACCGGCACTGGCCCGCGGATTCATGAATCCAGAGGCGCTGTGAAGCCTC 241
Qy 239 CAGCAACAAGATTCCTATCTTTTGGCCACCAATGAATGCAATTCCTAGGTCCCA 298

Db 242 CAGCAACAAGTATCCTATCTCTTTTGGCCACCCATGAATGCAATTCCTAGGTCCCA 301
Qy 299 AAGACCTTTTTCATATATAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACGGAAG 358
Db 302 AAGACCTTTTTCATATATAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACGGAAG 361
Qy 359 GATTTAACGAAGGATTTGTGGAAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTACC 418
Db 362 GATTTAACGAAGGATTTGTGGAAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTACC 421
Qy 419 AGGCAATTCAGCAACAGAGCTCTTCAGAACTCAGGGAGAGGTGGAATATCTGCAGATG 478
Db 422 AGGCAATTCAGCAACAGAGCTCTTCAGAACTCAGGGAGAGGTGGAATATCTGCAGATG 481
Qy 479 CAAGCAGTGAAGGAGGAGGTGATAGAGTGAAGAGTGAAGAGGCAAGAGAAAGATG 538
Db 482 CAAGCAGTGAAGGAGGAGGTGATAGAGTGAAGAGTGAAGAGGCAAGAGAAAGATG 541
Qy 539 AAAAAGCAGGCTCAAAACGGAAGAGTCAATATCTTCAAAGAAATCCTCTAAACAGTCCC 598
Db 542 AAAAAGCAGGCTCAAAACGGAAGAGTCAATATCTTCAAAGAAATCCTCTAAACAGTCCC 601
Qy 599 GGAATCTCAGAGATGAAGATGACAAAGACTGCAAAAGAGAGAAACAAAGAGAGCT 658
Db 602 GGAATCTCAGAGATGAAGATGACAAAGACTGCAAAAGAGAGAGAAACAAAGAGAGCT 661
Qy 659 CTGAGGCTGAGATCGGGCAAGCACAAGAAACAACTTCAGACTTCGAGAAACCA 718
Db 662 CTGAGGCTGAGATCGGGCAAGCACAAGAAACCACTTCAGACTTCGAGAAACCA 721
Qy 719 GTGAAGGAGCTTAACCTACCAATGAATGCTGATATTAAGAGAAACCAAGAGAGTTA 778
Db 722 GTGAAGGAGCTTAACCTACCAATGAATGCTGATATTAAGAGAAACCAAGAGAGTTA 781
Qy 779 TATGTTGGTGTCTAATATCTTGGATTTGATATGAACCAACATAGTCTCTGTTGTC 838
Db 782 TATGTTGGTGTCTAATATCTTGGATTTGATATGAACCAACATAGTCTCTGTTGTC 841
Qy 839 ATTGACAGAACCCAGTTGATGATATTCATATTCCTCTCTGTTGTTGCGGG 898
Db 842 ATTGACAGAACCCAGTTGATGATATTCATATTCCTCTCTGTTGTTGCGGG 901
Qy 899 GGAAGAGACATTTAGCCTTTTAAAGTTACTGATTTAAATTTCAATGTTTATTTGTTGC 958
Db 902 GGAAGAGACATTTAGCCTTTTAAAGTTACTGATTTAAATTTCAATGTTTATTTGTTGC 961
Qy 959 ATGAAGTCCCTTAACCACTAAGAGTATCAAGATTTTGGCGAGACTTATACATGCT 1018
Db 962 ATGAAGTCCCTTAACCACTAAGAGTATCAAGATTTTGGCGAGACTTATACATGCT 1021
Qy 1019 AGGATC 1024
Db 1022 AGGATC 1027

RESULT 8
AR220855
LOCUS AR220855
DEFINITION Sequence 96 from patent US 6426186.
ACCESSION AR220855
VERSION AR220855.1 GI:23227732
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4487)
AUTHORS Jones,K.A., Volkmuth,W. and Walker,M.G.
TITLE Bone remodeling genes
JOURNAL Patent: US 6426186-A 96 30-JUL-2002;
FEATURES Location/Qualifiers
1..4487
/organism="unknown"
BASE COUNT 1457 a 826 c 887 g 1314 t 3 others

Db 482 CAAGCAGTGTAGGAAGAGGTGATAGAGTAGAAGAGATGGAAGAGGCAAGAAAGAAATG 541
Qy 539 AAAAAGCAGGCTCAAAAACGGAAAAAGTCATATACCTCAAAAGAAATCCTCTAAACAGTCCC 598
Db 542 AAAAAGCAGGCTCAAAAACGGAAAAAGTCATATACCTCAAAAGAAATCCTCTAAACAGTCCC 601
Qy 599 GGAATCTCAGGAGATGAAGATGACAAAGCTGCAAAAGAGAGGAAAAAACAAGACAGCT 658
Db 602 GGAATCTCAGGAGATGAAGATGACAAAGCTGCAAAAGAGAGGAAAAAACAAGACAGCT 661
Qy 659 CTGAGGCTGAGATGCGGCAACGACACAAAGAACACAACTTCAGACTTCGAGAACCA 718
Db 662 CTGAGGCTGAGATGCGGCAACGACACAAAGAACACAACTTCAGACTTCGAGAACCA 721
Qy 719 GTGAAGGAGCTAACTACCAATGAATGCTGCATATTAAGAGAAAAACCAAA-GAAGGTT 777
Db 722 GTGAAGGAGCTAACTACCAATGAATGCTGCATATTAAGAGAAAAACCAAGAGGTT 781
Qy 778 ATATGTTGTTGTTCTAATATCTTGGATTGATATGAACCAACACATAGTCTCTTGTGT 837
Db 782 ATATGTTGTTGTTCTAATATCTTGGATTGATATGAACCAACACATAGTCTCTTGTGT 841
Qy 838 CATTCAGAACCCAGTTTGTATGATACATATTCATATTCCTCTCTGTTGTGTTTGGG 897
Db 842 CATTCAGAACCCAGTTTGTATGATACATATTCATATTCCTCTCTGTTGTGTTTGGG 901
Qy 898 GGGAAAAGACATTTAGCCCTTTTAAAGTTACTGATTTAAATTTCAATGTTTGGTTG 957
Db 902 GGGAAAAGACATTTAGCCCTTTTAAAGTTACTGATTTAAATTTCAATGTTTGGTTG 961
Qy 958 CATGAAGTTGCCCTTAACCACTAAGGATTATCAAGATTTTGGCGAGCTTATACATGTC 1017
Db 962 CATGAAGTTGCCCTTAACCACTAAGGATTATCAAGATTTTGGCGAGCTTATACATGTC 1021
Qy 1018 TAGGATC 1024
Db 1022 TAGGATC 1028

RESULT 7

ABT06906

ID ABT06906 standard; cDNA; 1664 BP.

XX

AC ABT06906;

XX 07-NOV-2002 (first entry)

XX

XX

XX Human ovarian cancer associated coding sequence SEQ ID NO: 68.

XX Human; ovarian cancer; cancer; gene; ss.

XX Homo sapiens.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PT Composition for detecting and treating ovarian cancer, comprises a
PT specific polypeptide, polynucleotide, T cell population, or antigen
PT presenting cell -
XX
XX Example 1; Fig 16; 188pp; English.
XX
CC The present invention relates to a method of detecting the presence of
CC ovarian cancer in a patient, involving detecting ovarian cancer
CC associated polynucleotides. The method is not only used to detect the
CC presence of cancer, preferably ovarian cancer in a patient, but also is
CC used to stimulate and/or expand T cells specific for an ovarian tumour
CC protein. The sequences can be used in vaccines used to treat cancer. The
CC present sequence is an ovarian cancer associated coding sequence.
XX
SQ Sequence 1664 BP; 566 A; 280 C; 319 G; 499 T; 0 other;

Query Match 80.9%; Score 828.4; DB 24; Length 1664;

Best Local Similarity 98.7%; Pred. No. 3.6e-311;

Matches 835; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

179 AGGGCTACCCGCACTGGCCGCGCGGATTTGATGAACCTCCAGAGGCGCTGTGAAGCCTC 238

17 AGGGCCGCGCAGTGCAGGCTAGCCCGATTTGATGAACCTCCAGAGGCGCTGTGAAGCCTC 76

239 CAGCAAAACAAGTATCTATCTTTTGGCAACCATGAACCTGCAATTTCTAGGTCCCA 298

77 CAGCAAAACAAGTATCTATCTTTTGGCAACCATGAACCTGCAATTTCTAGGTCCCA 136

299 AGACCTTTTTCATATAGGAGTACAAAGACAGTTTGGAAAGTCAAAACGGAAG 358

137 AGACCTTTTTCATATAGGAGTACAAAGACAGTTTGGAAAGTCAAAACGGAAG 196

359 GATTTAAACGAAGATTGTGGAAATAGAAATAACCCAGAGTAAAGTTTACTGGCTACC 418

197 GATTTAAACGAAGATTGTGGAAATAGAAATAACCCAGAGTAAAGTTTACTGGCTACC 256

419 AGGCAATTCAGCAACAGAGCTCTTTCAGAAACTGAGGAGAGGTGGAATATCTGCAGATG 478

257 AGGCAATTCAGCAACAGAGCTCTTTCAGAAACTGAGGAGAGGTGGAATATCTGCAGATG 316

479 CAAGCAGTGAAGAGAGGTGATAGAGTAGAAGAGATGGAAGAGGCAAAAGAAAGATG 538

317 CAAGCAGTGAAGAGAGGTGATAGAGTAGAAGAGATGGAAGAGGCAAAAGAAAGATG 376

539 AAAAAGCAGGCTCAAAACGGAAGAGTCAATATCTTCAAGAAATCCTCTAAACAGTCCC 598

377 AAAAAGCAGGCTCAAAACGGAAGAGTCAATATCTTCAAGAAATCCTCTTAAACAGTCCC 436

599 GGAATCTCCAGGAGATGAAGATGACAAAGACTGCAAGAGAGGAAAAAACAAGCAGCT 658

437 GGAATCTCCAGGAGATGAAGATGACAAAGACTGCAAGAGAGGAAAAAACAAGCAGCT 496

659 CTGAGGCTGAGATGCGGCAACGACACAAAGAAACAACTTCAGACTTCGAGAAACCA 718

497 CTGAGGCTGAGATGCGGCAACGACACAAAGAAACAACTTCAGACTTCGAGAAACCA 556

719 GTGAAGGAGCTAACTACCATATGATGCTGCATATTAAGAGAAACCAAGAGGTTA 778

557 GTGAAGGAGCTAACTACCATATGATGCTGCATATTAAGAGAAACCAAGAGGTTA 616

779 TATGTTTGGTTGTTCTAATATTTCTTGGATTGATATGAACCAACACATAGTCTTGTGTC 838

617 TATGTTTGGTTGTTCTAATATTTCTTGGATTGATATGAACCAACACATAGTCTTGTGTC 676

839 ATTGACAGAAACCCAGTTTGTATGATCATTAATTCATATTCCTCTCTGTTGTTTGGGG 898

677 ATTGACAGAAACCCAGTTTGTATGATCATTAATTCATATTCCTCTCTGTTGTTTGGGG 736

899 GGAAGAGACATTTAGCCCTTTTAAAGCTTAAAGTAAATTTTCAATGTTTATTTGCTGC 958

737 GGAAGAGACATTTAGCCCTTTTAAAGCTTAAAGTAAATTTTCAATGTTTATTTGCTGC 796

959 ATGAAGTTGCTTAAACCACTAAGGATTATCAAGATTTTGGCGAGACTTATACATGCT 1018

Db 797 ATGAAGTTCCTTACCACTAAGGATTATCAAGATTTTGGCGAGACTTATACATGTCT 856
QY 1019 AGGATC 1024
Db 857 AGGATC 862

RESULT 8
ID ABT07105 standard; cDNA; 1664 BP.
XX
AC ABT07105;
XX
DT 07-NOV-2002 (first entry)
XX
DE Human ovarian cancer associated coding sequence SEQ ID NO: 267.
XX
KW Human; ovarian cancer; cancer; gene; ss.
XX
OS Homo sapiens.
XX
XX US2002076715-A1.
XX
XX 20-JUN-2002.
XX
XX 06-JUN-2001; 2001US-0876889.
XX
XX 23-SEP-1998; 98US-0159320.
PR 08-FEB-1999; 99US-0246429.
PR 16-SEP-1999; 99US-0397787.
XX
XX (BENS/) BENSON D R.
PA (LODE/) LODES M J.
PA (MITC/) MITCHAM J L.
PA (KING/) KING G E.
XX
XX Benson DR, Lodes MJ, Mitcham JL, King GE;
XX WPI; 2002-598720/64.
XX
XX Composition for detecting and treating ovarian cancer, comprises a
PT specific polypeptide, polynucleotide, T cell population, or antigen
PT presenting cell -
XX
XX Example 1; Page 100-101; 188pp; English.
XX
XX The present invention relates to a method of detecting the presence of
CC ovarian cancer in a patient, involving detecting ovarian cancer
CC associated polynucleotides. The method is not only used to detect the
CC presence of cancer, preferably ovarian cancer in a patient, but also is
CC used to stimulate and/or expand T cells specific for an ovarian tumour
CC protein. The sequences can be used in vaccines used to treat cancer. The
CC present sequence is an ovarian cancer associated coding sequence.
XX
XX Sequence 1664 BP; 566 A; 280 C; 319 G; 499 T; 0 other;
SQ

Query Match 80.9%; Score 828.4; DB 24; Length 1664;
Best Local Similarity 98.7%; Pred. No. 3.6e-211;
Matches 835; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 179 AGGGCTACCGCACTGGCGCGCGGATGATGAACTCCAGAGGCGCTGTGAAGCCTC 238
Db 17 AGGGCGCGCGAGTGGCGGATGATGAACTCCAGAGGCGCTGTGAAGCCTC 76
QY 239 CAGCAAAACAAAGTATCTCTCTTTTGGCAACCATGAACTGCAATTTCTAGGTCCCA 298
Db 77 CAGCAAAACAAAGTATCTCTCTTTTGGCAACCATGAACTGCAATTTCTAGGTCCCA 136
QY 299 AAGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACCAACGGAAG 358
Db 137 AAGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACCAACGGAAG 196

QY 359 GATTTAACGAAGGATTTGGGAAATAGAAAATAAACCAGGAGTAAAGTTTACTGGCTACC 418
Db 197 GATTTAACGAAGGATTTGGGAAATAGAAAATAAACCAGGAGTAAAGTTTACTGGCTACC 256
QY 419 AGGCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGAGAAAGTGGAAATATCTGCAGATG 478
Db 257 AGGCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGAGAAAGTGGAAATATCTGCAGATG 316
QY 479 CAAGCAGTGAAGAGAGTGTATAGTAGAAGAGATGGAAGAAAGGCAAGAAAGAAATG 538
Db 317 CAAGCAGTGAAGAGAGTGTATAGTAGAAGAGATGGAAGAAAGGCAAGAAAGAAATG 376
QY 539 AAAAAGCAGGCTCAAAACGGAAGAGTCAATATCTCTCAAGAAATCTCTTAAACAGTCCC 598
Db 377 AAAAAGCAGGCTCAAAACGGAAGAGTCAATATCTCTCAAGAAATCTCTTAAACAGTCCC 436
QY 599 GGAATCTCCAGGAGATGAAGATGCAAGAGTCTGCAAGAGAGAGAAACCAAAACAGCT 658
Db 437 GGAATCTCCAGGAGATGAAGATGCAAGAGTCTGCAAGAGAGAGAAACCAAAACAGCT 496
QY 659 CTGAGGTTGGAGATGCGGGCAACGACACAAGAAACACAACCTTCAGACTTGCAGAAACCA 718
Db 497 CTGAGGTTGGAGATGCGGGCAACGACACAAGAAACACAACCTTCAGACTTGCAGAAACCA 556
QY 719 GTGAAGGACCTTAACCTACATAATGAATGCTGCATATATTAAAGAGAAACCAAGAAAGTTA 778
Db 557 GTGAAGGACCTTAACCTACATAATGAATGCTGCATATATTAAAGAGAAACCAAGAAAGTTA 616
QY 779 TATGTTGGTGTCTAATATTTCTTGATTTGATATGAACCAACATAGTCTCTGTTGTC 838
Db 617 TATGTTGGTGTCTAATATTTCTTGATTTGATATGAACCAACATAGTCTCTGTTGTC 676
QY 839 ATTGACAGAACCCAGTTTGTATGTACATATTCTCTCTCTGTTGTTGTTGCGGG 898
Db 677 ATTGACAGAACCCAGTTTGTATGTACATATTCTCTCTCTGTTGTTGTTGCGGG 736
QY 899 GGAAAGACATTTTAGCCTTTTAAAGTTTACTGATTAAATTTTCAATGTTTGGTTGC 958
Db 737 GGAAAGACATTTTAGCCTTTTAAAGTTTACTGATTAAATTTTCAATGTTTGGTTGC 796
QY 959 ATGAAGTTGCCCTTAACCACTAAGGATTTCAAGATTTTGGCGAGACTTATACATGCT 1018
Db 797 ATGAAGTTGCCCTTAACCACTAAGGATTTCAAGATTTTGGCGAGACTTATACATGCT 856
QY 1019 AGGATC 1024
Db 857 AGGATC 862
RESULT 9
ABX72784
ID ABX72784 standard; cDNA; 1664 BP.
XX
AC ABX72784;
XX
XX 14-MAR-2003 (first entry)
DT
XX
DE Human ovarian carcinoma antigen partial cDNA sequence #50.
XX
KW Human; cancer detection; ovarian carcinoma antigen; ovarian cancer;
KW tumour antigen; tumour; OV2; OV3; OV6; OV9; OV10; OV12; OV14; OV17;
KW OV18; OV23; OV24; OV27; OV41; OV54; OV57; gene; ss.
XX
OS Homo sapiens.
XX
XX US6468758-B1.
XX
XX 22-OCT-2002.
PD
XX
PF 16-SEP-1999; 99US-0397787.
XX
XX 23-SEP-1998; 98US-0159320.
PR 08-FEB-1999; 99US-0246429.
PR

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2004, 06:11:45 ; Search time 3909 Seconds
(without alignments)
10716.669 Million cell updates/sec

Title: US-09-787-328B-3
Perfect score: 1024
Sequence: 1 accgtctccgcgcggctt.....actatacatgtctaggatc 1024

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 2889711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sta.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sta.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1013	98.9	1423	9	AF110642	AF110642 Homo sapi
2	1013	98.9	1830	9	AF151900	AF151900 Homo sapi
3	1012	98.8	1973	6	BD156109	BD156109 Primer fo
4	1012	98.8	1973	9	AK001280	AK001280 Homo sapi
5	1001	97.8	1427	9	AB029156	AB029156 Homo sapi
6	990	96.7	1817	9	BC015483	BC015483 Homo sapi
7	980	95.7	3805	6	BD131023	BD131023 Human gco
8	969	94.6	4487	6	AR220855	AR220855 Sequence
9	828.4	80.9	1664	6	AR240312	AR240312 Sequence
10	828.4	80.9	1664	6	AR240311	AR240311 Sequence
11	772.8	75.5	2750	10	AF389347	AF389347 Rattus no
12	623.6	60.9	737	6	BD147622	BD147622 Primer fo
13	547.8	53.5	612	10	AB029493	AB029493 Mus muscu
14	425	41.5	1247	9	IR0783648	AL109689 Homo sapi
15	363	35.4	130005	9	AL590240	AL590240 Human DNA
16	299	29.2	2837	9	HSMB01382	AL133102 Homo sapi
17	299	29.2	148204	2	AC069400	AC069400 Homo sapi
18	299	29.2	157624	2	AC087666	AC087666 Homo sapi
19	299	29.2	182741	2	AC018910	AC018910 Homo sapi
20	299	29.2	184113	9	AC024370	AC024370 Homo sapi
21	263.6	25.7	334	6	BD054901	BD054901 Sequence
22	195	19.0	2096	9	BC018991	BC018991 Homo sapi
23	193.4	18.9	2376	6	AR081931	AR081931 Sequence
24	193.4	18.9	2376	6	AX302495	AX302495 Sequence
25	193.4	18.9	2376	6	E08546	E08546 DNA encodin
26	193.4	18.9	2376	9	HUMHDFG	D16431 Human mRNA
27	192.4	18.8	113199	9	AC103876	AC103876 Homo sapi
28	189	18.5	723	6	E08545	E08545 DNA encodin
29	188.6	18.4	869	6	AR081930	AR081930 Sequence
30	187	18.3	1932	4	BT4237996	AJ237996 Bos tauru
31	186	18.2	1752	10	AF251787	AF251787 Mus muscu
32	185	18.1	1744	10	BC005713	BC005713 Mus muscu
33	185	18.1	1763	10	BC021654	BC021654 Mus muscu
34	184.2	18.0	711	10	AF448810	AF448810 Rattus no
35	184.2	18.0	1571	10	AF389348	AF389348 Rattus no
36	183.8	17.9	714	6	E13960	E13960 Mouse mRNA
37	183.8	17.9	1563	6	E13961	E13961 Mouse mRNA
38	183.8	17.9	1563	10	MUSHDGF	D33707 Mus musculu
39	182.6	17.8	1966	10	AF339083	AF339083 Mus muscu
40	182.6	17.8	2559	10	AF339084	AF339084 Rattus no
41	182.6	17.8	2831	10	AF339082	AF339082 Mus muscu
42	181	17.7	965	10	BC002104	BC002104 Mus muscu
43	181	17.7	1255	10	MMU308966	AJ308966 Mus muscu
44	181	17.7	1853	10	BC052177	BC052177 Mus muscu
45	181	17.7	2928	10	MMU308965	AJ308965 Mus muscu

ALIGNMENTS

RESULT 1
AF110642
LOCUS AF110642 Homo sapiens hepatoma-derived growth factor 2 (HDGF2) mRNA, complete cds.
DEFINITION AF110642.1 GI:20530136
ACCESSION AF110642
VERSION AF110642.1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1423)
AUTHORS Yu,L., Fu,Q. and Tu,Q.
TITLE Cloning of a novel human cDNA which is a homolog to mouse

```

hepatoma-derived growth factor (mHDF) and termed hHDF2
Unpublished
2 (bases 1 to 1423)
Yu L., Zhang H.L., Fu, Q. and Zhao, Y.
Direct Submission
Submitted (04-DEC-1998) Lab of Human Gene Research, Institute of
Genetics, 220 Handan Road, Shanghai 200433, China
Location/Qualifiers
1. .1423
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. .1423
/gene="HDGF2"
125. .736
/gene="HDGF2"
/feature="similar to Mus musculus hepatoma-derived growth
factor"
/codon_start=1
/product="hepatoma-derived growth factor 2"
/protein_id="AA027001.1"
/db_xref="GI:20530137"
/translation="MARPPREYKAGDLVFAKMKGYPHWPARIDELPEGAVKPPANKY
PIFFGTHETAFGLPKDLFPYKEYDKFGSKNRKGFNEGLWEIENNPVGFYQAI
QQSSSTEGEGNTADASSEEDGVEEDGKGRKNEKAGSKKSYTSKSKSKQR
KSPGDEDDKDCKEENKSSSEGGDAGNDTRNTTSDLQKTSEGT"
polyA_signal
1393. .1398
/gene="HDGF2"
BASE COUNT 464 a 282 c 320 g 357 t
ORIGIN
Query Match 98.9%; Score 1013; DB 9; Length 1423;
Best Local Similarity 99.9%; Pred. No. 1.1e-237;
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 1 ACCGCTGTCGCGCGGCTTGAGCCCGCGGGGA-GGCGCGCAATTGTCGCGCCCGG 59
Db |||||
Qy 4 ACCGCTGTCGCGCGGCTTGAGCCCGCGGGAGCGCGCAATTGTCGCGCCCGG 63
Db |||||
Qy 60 GGGGCGGCTCCCGGCATCTCCGGCGGACCAAGACTACAGAGGGGGGGCTGG 119
Db |||||
Qy 64 GGGGCGGCTCCCGGCATCTCCGGCGGACCAAGACTACAGAGGGGGGGCTGG 123
Db |||||
Qy 120 GATGCGGCTCCGGGCGGCGGAGTCAAAAGCGGCGACCTGCTTCGCCAAGATGA 179
Db |||||
Qy 124 GATGCGGCTCCGGGCGGCGGAGTCAAAAGCGGCGACCTGCTTCGCCAAGATGA 183
Db |||||
Qy 180 GGGCTACCGCACTGGCGCGCGGATGATGAATCCAGAGGGCGCTGTGAAGCTCC 239
Db |||||
Qy 184 GGGCTACCGCACTGGCGCGCGGATGATGAATCCAGAGGGCGCTGTGAAGCTCC 243
Db |||||
Qy 240 AGCAACAAGATGATCTATCTCTTTTGGCAACCATGCAATCTAGTGTCCAA 299
Db |||||
Qy 244 AGCAACAAGATGATCTATCTCTTTTGGCAACCATGCAATCTAGTGTCCAA 303
Db |||||
Qy 300 AGACTTTTCCATATAGAGTCAAGAGCAAGATTTGGAAAGTCAACAAACGAAGG 359
Db |||||
Qy 304 AGACTTTTCCATATAGAGTCAAGAGCAAGATTTGGAAAGTCAACAAACGAAGG 363
Db |||||
Qy 360 ATTTAACAAGATGTTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGCTAC 419
Db |||||
Qy 364 ATTTAACAAGATGTTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGCTAC 423
Db |||||
Qy 420 GGCATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTTGAATCTCAGATGC 479
Db |||||
Qy 424 GGCATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTTGAATCTCAGATGC 483
Db |||||
Qy 480 AAGCAGTTCAGCAACAGAGTGTAGATAGAAAGATGAGAAAGCAAAAGAAATGA 539
Db |||||
Qy 484 AAGCAGTTCAGCAACAGAGTGTAGATAGAAAGATGAGAAAGCAAAAGAAATGA 543
Db |||||
Qy 540 AAAAGAGGCTCAAAACGGAAGAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCC 599
Db |||||

```

544 AAAAGCAGGCTCAAAACGGAAGAGTATATCTTCAAGAAATCTCTTAAACAGTCCC 603

600 GAAATCTCCAGGAGATGAAGATGACAAAGAGTCAAGAGAGAGGAAACAAACAGCTC 659

604 GAAATCTCCAGGAGATGAAGATGACAAAGAGTCAAGAGAGAGGAAACAAACAGCTC 663

660 TGAGGCTGAGATGCGGGCAACGACACAAAGAAACACAACTTCAGACTTGCAGAAAC 719

664 TGAGGCTGAGATGCGGGCAACGACACAAAGAAACACAACTTCAGACTTGCAGAAAC 723

720 TGAAGGACCTTAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779

724 TGAAGGACCTTAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 783

780 ATGTTGCTGCTTAATATCTTGGATTCATGATGATGATGATGATGATGATGATGAT 839

784 ATGTTGCTGCTTAATATCTTGGATTCATGATGATGATGATGATGATGATGATGAT 843

840 TTGACAGAACCCAGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 899

844 TTGACAGAACCCAGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 903

900 GAAAGACATTTAGCCCTTTTAAAGTTACGATTAATTTTATGATGATGATGATGATGAT 959

904 GAAAGACATTTAGCCCTTTTAAAGTTACTGATTAATTTTATGATGATGATGATGAT 963

960 TGAAGTTGCCCTTAACCACTAAGGATTCATCAAGATTTTTCGCGAGCTTATACATGCTA 1019

964 TGAAGTTGCCCTTAACCACTAAGGATTCATCAAGATTTTTCGCGAGCTTATACATGCTA 1023

1020 GGATC 1024

1024 GGATC 1028

RESULT 2

AF151900 1830 bp mRNA linear PRI 18-MAY-2000

LOCUS Homo sapiens CGI-142 protein mRNA, complete cds.

DEFINITION AF151900

ACCESSION AF151900

VERSION AF151900.1 GI:4929752

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1830)

AUTHORS Lai, C.H., Chou, C.Y., Ch'ang, I.Y., Liu, C.S. and Lin, W.

TITLE Identification of novel human genes evolutionarily conserved in

Caenorhabditis elegans by comparative proteomics

JOURNAL Genome Res. 10 (5), 703-713 (2000)

MEDLINE 20272150

PUBMED 10810093

REFERENCE 2 (bases 1 to 1830)

AUTHORS Lin, W.-C.

TITLE Direct Submission

JOURNAL Submitted (17-MAY-1999) Institute of Biomedical Sciences, Academia

Sinica, No. 128, Sec. II, Academia Road, Taipei 115, Taiwan

FEATURES

Location/Qualifiers

1. .1830

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="15"

125. .736

/codon_start=1

/product="CGI-142 protein"

/protein_id="A034137.1"

/db_xref="GI:4929753"

/translation="MARPPREYKAGDLVFAKMKGYPHWPARIDELPEGAVKPPANKY

PIFFGTHETAFGLPKDLFPYKEYDKFGSKNRKGFNEGLWEIENNPVGFYQAI

QQSSSTEGEGNTADASSEEDGVEEDGKGRKNEKAGSKKSYTSKSKSKQR

KSPGDEDDKDCKEENKSSSEGGDAGNDTRNTTSDLQKTSEGT"

CDS

BASE COUNT	588 a	342 c	385 g	515 t	
ORIGIN					
Query Match	98.9%; Score 1013; DB 9; Length 1830;				
Best Local Similarity	99.9%; Pred. No. 1,1e-237;				
Matches 1024; Conservative	0; Mismatches 0; Indels 1; Gaps 1;				
Qy	1	ACCGCTCGTCCGCGCGCTTTAGCGCCGCGGGGAGCGG-CGCAATTCGTGCGCCCGCGG	59		
Db	4	ACCGCTCGTCCGCGCGCTTTAGCGCCGCGGGAGCGCGCAATTCGTGCGCCCGCGG	63		
Qy	60	GGGGGGCGCTCCCGGCACTTCGCGCGGACCAAGGACTACCAAGGAGGGAGGGCGCTGG	119		
Db	64	GGGGGGCGCTCCCGGCACTTCGCGCGGACCAAGGACTACCAAGGAGGGAGGGCGCTGG	123		
Qy	120	GATGGCGCGTCCGCGCGCGCGGAGTACAAAGCGGGCGACTGGTCTTCGCCAAGATGAA	179		
Db	124	GATGGCGCGTCCGCGCGCGCGGAGTACAAAGCGGGCGACTGGTCTTCGCCAAGATGAA	183		
Qy	180	GGGCTACCCGCACTGGCGCGCGCGGATTTGATGAATCCCAAGAGGGCGCTGTGAAGCTCC	239		
Db	184	GGGCTACCCGCACTGGCGCGCGCGGATTTGATGAATCCCAAGAGGGCGCTGTGAAGCTCC	243		
Qy	240	AGCAACAACTATCTTCTCTTTTGGCACCCTGAAACTGCAATTTCTAGGTCCCAA	299		
Db	244	AGCAACAACTATCTTCTCTTTTGGCACCCTGAAACTGCAATTTCTAGGTCCCAA	303		
Qy	300	AGACCTTTTCCATATAAGGAGTACAAGACAAGTTTGGAAAGTCAACCAACCGAAGG	359		
Db	304	AGACCTTTTCCATATAAGGAGTACAAGACAAGTTTGGAAAGTCAACCAACCGAAGG	363		
Qy	360	ATTTAACGAAGGATTTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACCA	419		
Db	364	ATTTAACGAAGGATTTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACCA	423		
Qy	420	GGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGAGGAGGTTGGAATTACTGCAGATGC	479		
Db	424	GGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGAGGAGGTTGGAATTACTGCAGATGC	483		
Qy	480	AAGCAGTGAAGAGAGTGTATAGATGAGAGAGATGGAAGGCAAGGCAAGGAAAGATGA	539		
Db	484	AAGCAGTGAAGAGAGTGTATAGATGAGAGAGATGGAAGGCAAGGCAAGGAAAGATGA	543		
Qy	540	AAAAGCAGGCTCAAAACGGAAAAAGTGCATATCTTCAAAGAAATCTCTTAAACAGTCCCG	599		
Db	544	AAAAGCAGGCTCAAAACGGAAAAAGTGCATATCTTCAAAGAAATCTCTTAAACAGTCCCG	603		
Qy	600	GAATCTCCAGGAGATGAAGATGACAAGACTGCAAGAGAGAGAAACAAAAGCAGCTC	659		
Db	604	GAATCTCCAGGAGATGAAGATGACAAGACTGCAAGAGAGAGAAACAAAAGCAGCTC	663		
Qy	660	TGAGGTTGGAGATGCGGGCAACGACACAAAGAACTTCAGACTTCGAGAAACCCAG	719		
Db	664	TGAGGTTGGAGATGCGGGCAACGACACAAAGAACTTCAGACTTCGAGAAACCCAG	723		
Qy	720	TGAAGGCACTAACTACATAATGAATGCTGTCATATTAAGAGAAACCAACAAGAGTTAT	779		
Db	724	TGAAGGCACTAACTACATAATGAATGCTGTCATATTAAGAGAAACCAACAAGAGTTAT	783		
Qy	780	ATGTTGGTTGCTATATTTCTTGGATTGATATGAACCAACATAGTCTTGTGTGCA	839		
Db	784	ATGTTGGTTGCTATATTTCTTGGATTGATATGAACCAACATAGTCTTGTGTGCA	843		
Qy	840	TTGACAGAACCCAGATTGTATGATATTTCTATTTCTCTCTGTGTGTTTCGGGG	899		
Db	844	TTGACAGAACCCAGATTGTATGATATTTCTATTTCTCTCTGTGTGTTTCGGGG	903		
Qy	900	GAAGAAGCACTTTTAGCCCTTTTAAAGTACTGATTTAATTTTCATGTTATTTGGTTGCA	959		
Db	904	GAAGAAGCACTTTTAGCCCTTTTAAAGTACTGATTTAATTTTCATGTTATTTGGTTGCA	963		
Qy	960	TGAAGTTGCGCTTAAACCACTAAGGATTAACAAGATTTTTCGCGAGACTTATACATGTCTA	1019		
Db	964	TGAAGTTGCGCTTAAACCACTAAGGATTAACAAGATTTTTCGCGAGACTTATACATGTCTA	1023		
Qy	1020	GGATC 1024			
Db	1024	GGATC 1028			
RESULT 3					
LOCUS	BD156109	1973 bp	DNA	linear	PAT 17-JAN-2003
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.				
ACCESSION	BD156109	GI:27861867			
VERSION	JP 2002191363-A/10952				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1973)				
AUTHORS	Ota, T., Isogai, T., Nishikawa, T., Hayaashi, K., Saito, K., Yamamoto, J., Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.				
TITLE	Primer for synthesizing full-length cDNA and use thereof				
JOURNAL	Patent: JP 2002191363-A 10952 09-JUL-2002;				
COMMENT	HELIX RESEARCH INSTITUTE OS Homo sapiens (human) PN JP 2002191363-A/10952 PD 09-JUL-2002 PF 28-JUL-2000 JP 2000280990 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU, PI KEIICHI NAGAI, TETSUJI OTSUKI PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof PH Key Location/Qualifiers (156)..(764).				
FT	CDS	Location/Qualifiers			
source	1..1973	/organism="Homo sapiens"			
		/mol_type="genomic DNA"			
		/db_xref="taxon:9606"			
BASE COUNT	627 a	367 c	420 g	559 t	
ORIGIN					
Query Match	98.8%; Score 1012; DB 6; Length 1973;				
Best Local Similarity	99.9%; Pred. No. 2e-237;				
Matches 1023; Conservative	0; Mismatches 0; Indels 1; Gaps 1;				
Qy	2	CCGCTGTCGCGCGCGCTTTGAGCGCCGCGGGAGCGG-CGGCAATTCGTGCGCCCGCGG	60		
Db	36	CCGCTGTCGCGCGCGCTTTGAGCGCCGCGGGAGCGCGCGCAATTCGTGCGCCCGCGG	95		
Qy	61	GGGGGGCGCTCCCGGCACTTCGCGCGGACCAAGGACTACCAAGAGGGAGGGCGGTGGG	120		
Db	96	GGGGGGCGCTCCCGGCACTTCGCGCGGACCAAGGACTACCAAGAGGGAGGGCGGTGGG	155		
Qy	121	ATGGCGCGTCCGCGCGCGCGCGAGTACAAAGCGGGCGACCTGGTCTTCGCCAAGATGAAG	180		
Db	156	ATGGCGCGTCCGCGCGCGCGCGAGTACAAAGCGGGCGACCTGGTCTTCGCCAAGATGAAG	215		
Qy	181	GGCTACCCGCACTGGCGCGCGCGGATTTGATGAATCCCGAGAGGGCGCTGTGAAGCTCCA	240		
Db	216	GGCTACCCGCACTGGCGCGCGCGGATTTGATGAATCCCGAGAGGGCGCTGTGAAGCTCCA	275		
Qy	241	GCAACAACTATCTCTATCTTTTGGCACCCATGAACTGCAATTTCTAGGTCCCAA	300		
Db	276	GCAACAACTATCTCTATCTTTTGGCACCCATGAACTGCAATTTCTAGGTCCCAA	335		
Qy	301	GACCTTTTCCATATTAAGGAGTACAAAGCAAGTTTGGAAAGTCAACAAACCGGAAGGA	360		

Db 336 GACCTTTTTCATATAAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACAAACGGAAGGA 395
Qy 361 TTTACGAGGATTTGGGAATAGAAATATACCCAGGAGTAAAGTTTACTTGGCTACCAG 420
Db 396 TTTACGAGGATTTGGGAATAGAAATATACCCAGGAGTAAAGTTTACTTGGCTACCAG 455
Qy 421 GCAATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAAAGTGGAAATACTGCAGATGCA 480
Db 456 GCAATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAAAGTGGAAATACTGCAGATGCA 515
Qy 481 AGCAGTGAAGAAAGGATGATAGAGTAGAAGAGTGAAGAAAGGCAAAAGAAAGAAATGAA 540
Db 516 AGCAGTGAAGAAAGGATGATAGAGTAGAAGAGTGAAGAAAGGCAAAAGAAAGAAATGAA 575
Qy 541 AAGCAGGCTCAAAACGGAAGATCATATCTTCAAGAAATCTCTAAACAGTCCCGG 600
Db 576 AAGCAGGCTCAAAACGGAAGATCATATCTTCAAGAAATCTCTAAACAGTCCCGG 635
Qy 601 AATCTCCAGGAGATGAAGTACAAAGACTCAAAAGAGAGAGAAACAAAGACAGCTCT 660
Db 636 AATCTCCAGGAGATGAAGTACAAAGACTCAAAAGAGAGAGAAACAAAGACAGCTCT 695
Qy 661 GAGGTGAGATGCGGGCAACACACAGAAACACAACTTCAGACTTGCAGAAAACCAAGT 720
Db 696 GAGGTGAGATGCGGGCAACACACAGAAACACAACTTCAGACTTGCAGAAAACCAAGT 755
Qy 721 GAAGGACCTTAACCTAATGAATGCTGCATATTAAGAGAAACCAAGAGGTTATA 780
Db 756 GAAGGACCTTAACCTAATGAATGCTGCATATTAAGAGAAACCAAGAGGTTATA 815
Qy 781 TGTTGGTGTCTAATATCTTGGATTTGATATGAACCAACATAGTCTTGTGTCTAT 840
Db 816 TGTTGGTGTCTAATATCTTGGATTTGATATGAACCAACATAGTCTTGTGTCTAT 875
Qy 841 TGACAGAACCCAGTTGATGATACATATTCATATTCCTCTCTGTGTGTGTGTCGGGG 900
Db 876 TGACAGAACCCAGTTGATGATACATATTCATATTCCTCTCTGTGTGTGTGTCGGGG 935
Qy 901 AAAACACATTTAGCTTTTAAAGTTACTGATTAATTCATGTTATTTGTTGAT 960
Db 936 AAAACACATTTAGCTTTTAAAGTTACTGATTAATTCATGTTATTTGTTGAT 995
Qy 961 GAAGTTGCCCTTAAACACATGAAGATTAACAGATTTTTCGAGACTTATACATGCTAG 1020
Db 996 GAAGTTGCCCTTAAACACATGAAGATTAACAGATTTTTCGAGACTTATACATGCTAG 1055
Qy 1021 GATC 1024
Db 1056 GATC 1059

RESULT 4

AK001280

LOCUS

DEFINITION Homo sapiens cDNA FLJ10418 fis, clone NT2RP1000130, moderately similar to HEPATOMA-DERIVED GROWTH FACTOR.

ACCESSION AK001280

VERSION AK001280.1 GI:7022435

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS

Isogai, T., Ota, T., Hayaashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Wagatsuma, M., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.
NEDO human cDNA sequencing project
Unpublished

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

2 (bases 1 to 1973)
Isogai, T. and Otsuki, T.
Direct Submission
Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

FEATURES

source

1..1973
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP1000130"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP1"
/note="cloning vector: pUC19FL3-mRNA from NT2 neuronal precursor cells after 48-hours retinoic acid (RA) induction."
156..767
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA91597.1"
/db_xref="GI:7022436"
/translation="MARPRPYKAGDLVFAKMGYPHPARIDELPEGAVKPPANKYPIFPFGTHETAFGLPKDLFPYKEYKDFKSNKRGNEGLWEIENPVGVTGYQAIQQQSSSTEGEGNTADASSEEDRVEDGKGRKKEKAGSKKSKTSKSKQSRKSPDEDDKCKEENKSSSEGGAGNDTRNTTSDLOKTSBGT"

CDS

BASE COUNT 627 a 367 c 420 g 559 t
ORIGIN
Query Match 98.8%; Score 1012; DB 9; Length 1973;
Best Local Similarity 99.9%; Pred. No. 2e-237;
Matches 1023; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 2 CCGCTCGTCCCGCGCTTGAGGCGCGCGGAGCGC-GCGCAATTCGTGCGCGCGG 60
Db 36 CCGCTCGTCCCGCGCTTGAGGCGCGCGGAGCGCGCGCAATTCGTGCGCGCGG 95
Qy 61 GGGGCGGCTCCCGCATCTTCGCGCGACCAAGACTACCAAGAGGGAGCGGCTGG 120
Db 96 GGGGCGGCTCCCGCATCTTCGCGCGACCAAGACTACCAAGAGGGAGCGGCTGG 155
Qy 121 ATGGCGGCTCCCGCGCGCGAGTACAAAGCGGCGACCTGCTTCGCGCAAGATGAAG 180
Db 156 ATGGCGGCTCCCGCGCGCGAGTACAAAGCGGCGACCTGCTTCGCGCAAGATGAAG 215
Qy 181 GGCTACCGCGCTGCGCGCGCGGATGATGAATCCCGAGAGGCGCTGTGAAGCTCCA 240
Db 216 GGCTACCGCGCTGCGCGCGCGGATGATGAATCCCGAGAGGCGCTGTGAAGCTCCA 275
Qy 241 GCAAAACAGTATCTCTCTTTTGGGACCAATGAATGCAATTTCTAGTCCCAAA 300
Db 276 GCAAAACAGTATCTCTCTTTTGGGACCAATGAATGCAATTTCTAGTCCCAAA 335
Qy 301 GACCTTTTCCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAGGA 360
Db 336 GACCTTTTCCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAGGA 395
Qy 361 TTTAAACAGGATTTGGGAAATAGAAATAACCCAGAGTAAAGTTTACTTGGCTACCAG 420
Db 396 TTTAAACAGGATTTGGGAAATAGAAATAACCCAGAGTAAAGTTTACTTGGCTACCAG 455
Qy 421 GCAATTCAGCAACAGAGCTCTTCTGAAACTGAGGAGAGGTGGAAATACTGCAGATGCA 480
Db 456 GCAATTCAGCAACAGAGCTCTTCTGAAACTGAGGAGAGGTGGAAATACTGCAGATGCA 515

```
Qy 481 AGCAGTGAAGAGAGGTAGATAGATAGAGAGATGAAAGGCAAAAGAAAGAAATGAA 540
Db 516 AGCAGTGAAGAGAGGTAGATAGATAGAGAGATGAAAGGCAAAAGAAAGAAATGAA 575
Qy 541 AAACGAGGCTCAAAACGGAAAAAGTCAATATCTTCAAGAAATCTCTAAACAGTCCGG 600
Db 576 AAACGAGGCTCAAAACGGAAAAAGTCAATATCTTCAAGAAATCTCTAAACAGTCCGG 635
Qy 601 AAATCTCCAGAGATGAAGATGACAAAGACTGCAAAAGAGAGGAAAAAAGAGCAGTCT 660
Db 636 AAATCTCCAGAGATGAAGATGACAAAGACTGCAAAAGAGAGGAAAAAAGAGCAGTCT 695
Qy 661 GAGGCTGAGATGCGGCAACGACACAAAGAAACAATCTCAGACTTGCAGAAAAACAGT 720
Db 696 GAGGCTGAGATGCGGCAACGACACAAAGAAACAATCTCAGACTTGCAGAAAAACAGT 755
Qy 721 GAAGGACCTTAACCTACCATTAATGCTGCATATTAAGAGAACCAAGAGGTTATA 780
Db 756 GAAGGACCTTAACCTACCATTAATGCTGCATATTAAGAGAACCAAGAGGTTATA 815
Qy 781 TGTGTGTTGTCTAATATTTCTTGGATTGATATGAACCAACACATAGTCTCTTGTTCAT 840
Db 816 TGTGTGTTGTCTAATATTTCTTGGATTGATATGAACCAACACATAGTCTCTTGTTCAT 875
Qy 841 TGACAGAACCCAGTTTGTATGATATTAATTCATATTCCTCTCTGTGTGTTCGCGGG 900
Db 876 TGACAGAACCCAGTTTGTATGATATTAATTCATATTCCTCTCTGTGTGTTCGCGGG 935
Qy 901 AAAAGACATTTTACCCCTTTTAAAGTTACTGATTTAAATTCATGTTTGTTCAT 960
Db 936 AAAAGACATTTTACCCCTTTTAAAGTTACTGATTTAAATTCATGTTTGTTCAT 995
Qy 961 GAAGTGGCCCTTAACCACTAAGGATTATCAAGATTTTGGCAGACTTATACATGCTAG 1020
Db 996 GAAGTGGCCCTTAACCACTAAGGATTATCAAGATTTTGGCAGACTTATACATGCTAG 1055
Qy 1021 GATC 1024
Db 1056 GATC 1059
```

```
RESULT 5
AB029156 1427 bp mRNA linear PRI 19-FEB-2000
LOCUS Homo sapiens HRP-3 mRNA, complete cds.
DEFINITION
AB029156 Homo sapiens HRP-3 mRNA, complete cds.
ACCESSION
AB029156.1 GI:6855467
VERSION
AB029156.1
KEYWORDS
HRP-3.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (sites)
Ikegane, K., Yamamoto, M., Kishima, Y., Enomoto, H., Yoshida, K.,
Suemura, M., Kishimoto, T. and Nakamura, H.
A new member of a hepatoma-derived growth factor gene family can
translocate to the nucleus
Biochem. Biophys. Res. Commun. 266 (1), 81-87 (1999)
JOURNAL
MEDLINE
20050055
PUBMED
10581169
REFERENCE
2 (bases 1 to 1427)
Ikegane, K. and Nakamura, H.
Direct Submission
TITLE
Submitted (21-JUN-1999) Kazuhiro Ikegane, Osaka University Medical
School, Department of Molecular Medicine; 2-2 Yamada-oka, Suita,
Osaka 565-0871, Japan (E-mail:nakamura@med3.med.osaka-u.ac.jp,
Tel:81-6-6879-3837(ex.3837), Fax:81-6-6879-3839)
LOCATION/Qualifiers
Location/Qualifiers
1..1427
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="15"
```

```

/map="15q25"
/cell_type="lymphocytes"
1..1427
/gene="HRP-3"
/CD5
121..732
/gene="HRP-3"
/codon_start=1
/product="HRP-3"
/protein_id="BAA90477.1"
/db_xref="GI:6855468"
/translation="MARPRPREYKAGDLVPAWKGYPHWPARIDELPEGNVKEPANKY
PIFFGTHETAFGLPKDLFPYKDFKFSKNNKRGFNGLWEIENNPGVKFTGYQAI
QQSSSETTEGGNTADASBEEDRVBEGDKRKRKXKGRKSKYSYTSKSSKOSR
KSPGDDDDKDCKEENKSSSEGDGNDTRNTTSDLOKTSEGT"
BASE COUNT 454 a 287 c 327 g 359 t
ORIGIN
Query Match 97.8%; Score 1001; DB 9; Length 1427;
Best Local Similarity 99.8%; Pred. No. 9.4e-235;
Matches 1023; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
Qy 1 ACCGCTCGTCCGCCCGCTTGAGGCCCGCGGGAGCGC-GCGCAATTCGTCGCCCGCGG 59
Db 1 ACCGCTCGTCCG-CCGGCTTGAGGCCCGCGGGAGCGCGCAATTCGTCGCCCGCGG 59
Qy 60 GGGGGGGCGCTCCCGGCATCTTCGCGCGCAACCAAGGACTACCAAGAGGAGGAGCGGCTGG 119
Db 60 GGGGGGGCGCTCCCGGCATCTTCGCGCGCAACCAAGGACTACCAAGAGGAGGAGGCGGCTGG 119
Qy 120 GATGGCGCTGCGCGCCCGCGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 179
Db 120 GATGGCGCTGCGCGCCCGCGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 179
Qy 180 GGGCTACCGCCTGCGCGCCCGCGGATGATGAACTCCAGAGGCGGCTGTGAAGCTCC 239
Db 180 GGGCTACCGCCTGCGCGCCCGCGGATGATGAACTCCAGAGGCGGCTGTGAAGCTCC 239
Qy 240 AGCAAAACAAGTATCTTCTTTTGGCACCCTGAACTGCAATTCCTAGTCCCAA 299
Db 240 AGCAAAACAAGTATCTTCTTTTGGCACCCTGAACTGCAATTCCTAGTCCCAA 299
Qy 300 AGACCTTTTCCATATAAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACGGAAGG 359
Db 300 AGACCTTTTCCATATAAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACGGAAGG 359
Qy 360 ATTTAAGCAAGGATTTGTGGAAATAGAAATACCCAGGATTAAGTTTACTGGCTACCA 419
Db 360 ATTTAAGCAAGGATTTGTGGAAATAGAAATACCCAGGATTAAGTTTACTGGCTACCA 419
Qy 420 GGCAATTCAGCAACAGAGCTTTTCAGAACTGAGGAGAGAGGTGGAAATACTGCAGATGC 479
Db 420 GGCAATTCAGCAACAGAGCTTTTCAGAACTGAGGAGAGAGGTGGAAATACTGCAGATGC 479
Qy 480 AAGCAGTGAAGAAAGGTGATAGATAGAAAGATGGAAGGCAAAAGCAAAAGAAAGATGA 539
Db 480 AAGCAGTGAAGAAAGGTGATAGATAGAAAGATGGAAGGCAAAAGCAAAAGAAAGATGA 539
Qy 540 AAAAGCAGGCTCAAAACGGAAAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCCG 599
Db 540 AAAAGCAGGCTCAAAACGGAAAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCCG 599
Qy 600 GAAATCTCCAGGAGATGAAGATGACAAAGACTGCAAGAGAGGAAAAAAGAGCAGCTC 659
Db 600 GAAATCTCCAGGAGATGAAGATGACAAAGACTGCAAGAGAGGAAAAAAGAGCAGCTC 659
Qy 660 TGAAGGTGAGATGCGGGCAACGACACAAAGAAACAACACTTTCAGACTTGCAGAAAAACCG 719
Db 660 TGAAGGTGAGATGCGGGCAACGACACAAAGAAACAACACTTTCAGACTTGCAGAAAAACCG 719
Qy 720 TGAAGGGAACCTAATACCATTAATGAATGCTGCATATTAAGAGAAACCAAGAGGTTAT 779
Db 720 TGAAGGGAACCTAATACCATTAATGAATGCTGCATATTAAGAGAGAAACCAAGAGGTTAT 779
```

```
QY 780 ATGTTTGGTGTCTAATATTCTTGAGTTTGATATGAACCAACACATAGTCTTGTGTGCA 839
Db 780 ATGTTTGGTGTCTAATATTCTTGAGTTTGATATGAACCAACACATAGTCTTGTGTGCA 839
QY 840 TTGACAGAACCCAGTTTGATGTACATTTATTCATATTCCTCTCTGTTGTGTTGGGG 899
Db 840 TTGACAGAACCCAGTTTGATGTACATTTATTCATATTCCTCTCTGTTGTGTTGGGG 899
QY 900 GAAAGACACATTTTGTAGCTTTTAAAGTTTACTGATTTTAAATTTTCAATTTTGGTTGCA 959
Db 900 GAAAGACATTTTGTAGCTTTTAAAGTTTACTGATTTTAAATTTTCAATTTTGGTTGCA 959
QY 960 TGAAGTTGCCCTTAACCACTAAGGATTATCAAGATTTTTCGCGAGACTTATACATGTCTA 1019
Db 960 TGAAGTTGCCCTTAACCACTAAGGATTATCAAGATTTTTCGCGAGACTTATACATGTCTA 1019
QY 1020 GGATC 1024
Db 1020 GGATC 1024

RESULT 6
BC015483
LOCUS BC015483
DEFINITION Homo sapiens, CGI-142, clone MGC:8892 IMAGE:3913558, mRNA, complete cds.
ACCESSION BC015483
VERSION BC015483.1 GI:15930079
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1817)
Strausberg.R.
Direct Submission
Submitted (01-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: http://image.llnl.gov
Series: IRAK Plate: 22 Row: 1 Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7705319.
Location/Qualifiers
1. .1817
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:50810"
/db_xref="taxon:9606"
/clone="MGC:8892 IMAGE:3913558"
/tissue_type="Uterus, leiomyosarcoma"
/clone_lib="NIH MGC_71"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
99. .710
/codon_start=1
/product="CGI-142"
/protein_id="AAH15483.1"

CDS
FEATURES
source
```

```
/db_xref="GI:15930080"
/translation="MARPRREYKAGDILVFARMKGYPHWPARIDELPEGAVKPPANKY
PIFFGTHETAPLGPDLFPYKWKDFKGNKRGKGFHWELWENPVGFTGYQAI
QQSSSTFEGSGNTADASSSEEDRVEEDGKGRKNEKAGSKSKSYTKSSKQSR
KSPDEDDKCKEENKSSSGGSDAGNTRNTTSDLOKTSBGT"
BASE COUNT 601 a 329 c 376 g 511 t
ORIGIN
Query Match 96.7%; Score 990; DB 9; Length 1817;
Best Local Similarity 99.9%; Pred. No. 5e-232;
Matches 1001; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 24 GCCCGCGGGGAGCGC-GCGCAATTGTCGCGCCCGCGGGGGCGGCCCTCCCGGCATCTTC 82
Db 1 GCCCGCGGGGAGCGCGCAATTGTCGCGCCCGCGGGGGCGGCCCTCCCGGCATCTTC 60
QY 83 GCGGCGACCAAGGACTACCAAGAGGGGAGCGGCTGGGATGGCGGCTCGCGGCCCGCGG 142
Db 61 GCGGCGACCAAGGACTACCAAGAGGGGAGCGGCTGGGATGGCGGCTCGCGGCCCGCGG 120
QY 143 AGTACAAAGCGGGGAGCTGCTTCGCAAGATGAAGGGCTACCGCACTGGCGCGGCC 202
Db 121 AGTACAAAGCGGGGAGCTGCTTCGCAAGATGAAGGGCTACCGCACTGGCGCGGCC 180
QY 203 GGATTGATGAATCCCGAGAGGCGCTGTGAAGCCCTCCAGCAAAACAAGTATCCTATCTTC 262
Db 181 GGATTGATGAATCCCGAGAGGCGCTGTGAAGCCCTCCAGCAAAACAAGTATCCTATCTTC 240
QY 263 TTTTGGCACCCATGAAACTGCAATTTCTAGTTCCTGAGGCTTTTTCATATAGGAGT 322
Db 241 TTTTGGCACCCATGAAACTGCAATTTCTAGTTCCTGAGGCTTTTTCATATAGGAGT 300
QY 323 ACAAGACAAGTTTGGAAAGTCAACAAACGGAAGGATTTTAAAGCAAGGATTTGTCGGA 382
Db 301 ACAAGACAAGTTTGGAAAGTCAACAAACGGAAGGATTTTAAAGCAAGGATTTGTCGGA 360
QY 383 TAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACAGGCAATTCAGCAACAGAGCTCTT 442
Db 361 TAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACAGGCAATTCAGCAACAGAGCTCTT 420
QY 443 CAGAACTGAGGAGAGAGGTGGAATATCTGCAGATGCAAGCAGTGAGGAAGAGTGATA 502
Db 421 CAGAACTGAGGAGAGAGGTGGAATATCTGCAGATGCAAGCAGTGAGGAAGAGTGATA 480
QY 503 GAGTAGAAGAGAGTGGAAAGGCAAAAGAAAGATGAAAGCAGGCTCAAAAACGGA 562
Db 481 GAGTAGAAGAGAGTGGAAAGGCAAAAGAAAGATGAAAGCAGGCTCAAAAACGGA 540
QY 563 AGTCATATCTTCAAGAAATCCTCTAAACAGTCCCGGAAATCTCCAGGAGATGAAGTG 622
Db 541 AGTCATATCTTCAAGAAATCCTCTAAACAGTCCCGGAAATCTCCAGGAGATGAAGTG 600
QY 623 ACAAGACTGCAAGAGAGAGAAACAAAAGCAGCTCTGAGGGTGAGATGCGGGCAACG 682
Db 601 ACAAGACTGCAAGAGAGAGAAACAAAAGCAGCTCTGAGGGTGAGATGCGGGCAACG 660
QY 683 ACACAGAAACACACACTTCAGACTTGCAGAAACCCAGTGAAGGAGCCCTAACCATAT 742
Db 661 ACACAGAAACACACACTTCAGACTTGCAGAAACCCAGTGAAGGAGCCCTAACCATAT 720
QY 743 GAATGCTGCATATTAAAGAGAAACCAAGAGGTTATATGTTGGTTGCTTAATATCTT 802
Db 721 GAATGCTGCATATTAAAGAGAAACCAAGAGGTTATATGTTGGTTGCTTAATATCTT 780
QY 803 GGATTGATGAACCAACACATAGTCTTGTGTGTCATGTACAGAACCCCGAGTTGTATG 862
Db 781 GGATTGATGAACCAACACATAGTCTTGTGTGTCATGTACAGAACCCCGAGTTGTATG 840
QY 863 TACATTATTCATATTCCTCTCTGTTGTTTCGGGGGGAAGACATTTTACCCCTTTT 922
Db 841 TACATTATTCATATTCCTCTCTGTTGTTTCGGGGGGAAGACATTTTACCCCTTTT 900
QY 923 AAAAGTTACTGATTTAATTTTTCATGTTTATTTGTTGTCATGAAGTTGCGCCCTTAACCACTAAG 982
```


Db	4	CCGCTCGAGCGCCGGCTCGAGGCCCGGGGAGCGGCGCGTTCGTTCGTCCCGCGGA	63
Qy	61	GGGCGGCTCCCGGCATCTTCGGCGGACCAAGAGCTACCAAGAAAGGAGCGGCTGGG	120
Db	64	GGGCGGCTCCCGGCATCTTCGGCGGACCAAGAGCTACCAAGAAAGGAGCGGCTGGG	123
Qy	121	ATGGGGCGTCCGGGCCCCGGAGTACAAAGCGGGCACTGTGCTTCGCGCAGAGTGAAG	180
Db	124	ATGGCGGCTCCGGGCCCCGGAGTACAAAGCGGGCACTGTGCTTCGCGCAGAGTGAAG	183
Qy	181	GGCTACCGGCACTGGCGGGCCGGATTTGATGAATCCCGAGAGGCGCTGTGAAGCCCTCCA	240
Db	184	GGCTACCGGCACTGGCGGGCCGGATTTGATGAATCCCGAGAGGCGCTGTGAAGCCCTCCA	243
Qy	241	GCAAAAGATTCATCTCTCTTTTGGCACCCATGAATCGAATCTTAGGTCCTCAAA	300
Db	244	GCAAAAGATTCATCTCTCTTTTGGCACCCATGAATCGAATCTTAGGTCCTCAAA	303
Qy	301	GACCTTTTCCATATAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACAACGGAAGGA	360
Db	304	GACCTTTTCCATATAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACAACGGAAGGA	363
Qy	361	TTTAAACGAAGGATTTGGGAAATAGAAAATAAACCCAGGAGTAAAGTTTACTTGGCTACCA	420
Db	364	TTTAAACGAAGGATTTGGGAAATAGAAAATAAACCCAGGAGTAAAGTTTACTTGGCTACCA	423
Qy	421	GCAATTCAGCAACAGAGCTCTTCAGAAATCGAGGGAGAGGTGGAATATCTGCAGATGCA	480
Db	424	ACAATTCAGCAACAGAGCTCTTCAGAAATCGAGGGAGAGGTGGAATATCTGCAGATGCA	483
Qy	481	AGCAGTCAGGAAGAGGTGATAGAGTAGAAGATGGAAGGCAAAAGCAAAAGAAAGATGAA	540
Db	484	AGCAGTCAGGAAGAGGTGATAGAGTAGAAGATGGAAGGCAAAAGCAAAAGAAAGATGAA	540
Qy	541	AAAGCGCTCAAAACGGAAAGAGTCAATATCTTCAAAGAAATCCTCTAAACAGTCCCGG	600
Db	541	AAAGCGGCTCAAAACGGAAAGAGTCTCAATCTTCAAAGAAATCCTCTTAAACAGTCCCGG	600
Qy	601	AAATCTCCAGAGATGAAGATGACAAAGACTGCAAGAGAGGGAAGAAACAAGACGCTCT	660
Db	601	AAATCTCCAGAGAGCAAGATGACAAAGACTGCAAGAGAGGGAAGAAACAAGACGCTCT	660
Qy	661	GAGGTGGAGATCGGGCAACGACACAAGAAACAACATTCAGACTTCAGAAAAACAGT	720
Db	661	GAGGCGGAGATGCTGGCAATGACACAAGAAACAACGACTTCAGACTTCGAGAAAGCGGT	720
Qy	721	GAAGGACCTTAACCTACATATGATGCTGCTGATTAAGAGAAACCAACAAGAGTTATA	780
Db	721	GAAGGACCTTAACCTACGCTAATGATGCTGCTGATTAAGAGAAACCAACAAGAGTTAA	780
Qy	781	TGTTTGTGTCTAATATCTTGGATTTGATATGAACCAACACATAGTCTCTGTGTGCAT	840
Db	781	TGTTTGTGTCTGA---TTCTGATTTGATATGAACCAAC---AGTCTTTGTGTGCAT	833
Qy	841	TGACAGAACCCAGTTTGTATGTACATTAATTCATTAATTCATTAATTCATTAATTCAT	900
Db	834	TGACAAAGCCCGAGTGTGTGCATTT---ACATTTCTCTCTGTTCTGCTGAAGGAA	889
Qy	901	AAAAGACATTTTAGCTTTTAAAAGTACTGATTTAAATTCATTAATTCATTAATTCAT	960
Db	890	AAAAGACATTTTAACTTTTAAAGGTATTTGACTTAATTTATGTTATTTGTTGTCAT	949
Qy	961	GAAGTTCCTTAACCACTAAGGATTAATCAAGATTTTGTGGCAGACTTATACATGCTAG	1020
Db	950	GAGTTCCTTAACCACTAAGATGATCAAGATTTT---GCAGACTTCTCCGTGCTAG	1006
Qy	1021	GATC	1024
Db	1007	GCTC	1010

RESULT 12
BD147622

LOCUS	BD147622	737 bp	DNA	linear	PAT 17-JAN-2000
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.				
ACCESSION	BD147622				
VERSION	BD147622.1	GI:27853380			
KEYWORDS	JP 2002191363-A/2465				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 737)				
AUTHORS	Ota,T., Isogai,T., Nishikawa,T., Hayaashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.				
TITLE	Primer for synthesizing full-length cDNA and use thereof				
JOURNAL	Patent: JP 2002191363-A/2465 09-JUL-2002;				
COMMENT	HELIIX RESEARCH INSTITUTE OS Homo sapiens (human) PN JP 2002191363-A/2465 PD 09-JUL-2002 PF 28-JUL-2000 JP 2000280990 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SAITO, PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU, PI KEIICHI NAGAI, TETSUJI OTSUKI PC C12N15/09, C07K34/47, C07K36/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC 10, CC C12P21/02, C12P1/68/ C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key				
FEATURES	FT source	1..737			
source	Location/Qualifiers	/organism="Homo sapiens (human)";			
	1..737				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
BASE COUNT	227 a	180 c	204 g	117 t	9 others
ORIGIN					
Query Match	60.9%	Score 623.6;	DB 6;	Length 737;	
Best Local Similarity	96.2%	Pred. No. 3.8e-142;			
Matches	676;	Conservative	0;	Mismatches 23;	Indels 4; Gaps 4;
Qy	2	CCGCTCGTCGCGCCGCGCTTCAGGCGCGCGGGAGCGC-GGCAATTCGTCGGCCCGCGGG 60			
Db	36	CCGCTCGTCGCGCGCGCTTCAGGCGCGCGGGAGCGCAATTCGTCGGCCCGCGGG 95			
Qy	61	GGGCGCGCTCCCGGCATCTTCGCGCGACCAAGGACTACACAGAAAGGGAGCGGTGGG 120			
Db	96	GGGCGCGCTCCCGGCATCTTCGCGCGACCAAGGACTACACAGAAAGGGAGCGGTGGG 155			
Qy	121	ATGCGCGGTCGCGCGCCCGCGAGTACAAAGCGGGGCACTGTGCTTCGCCAAGATGAAG 180			
Db	156	ATGCGCGGTCGCGCGCCCGCGAGTACAAAGCGGGGCACTGTGCTTCGCCAAGATGAAG 215			
Qy	181	GGCTACCGGCATCGGCGCGCCGCGATTCGTAAGTGAAGTCCAGAGGCGCTGGAAGCTCCA 240			
Db	216	GGCTACCGGCATCGGCGCGCCGCGATTCGTAAGTGAAGTCCAGAGGCGCTGGAAGCTCCA 275			
Qy	241	GCAACCAAGATATCTATCTCTTTTGGCACCCATGAACTGCATTTCTAGGTCCCAA 300			
Db	276	GCAACCAAGATATCTATCTCTTTTGGCACCCATGAACTGCATTTCTAGGTCCCAA 335			
Qy	301	GACCTTTTCCATATAAGGAGTACAAAGCAAGTTGGAAAGTCAACAAACGGAAGGA 360			
Db	336	GACCTTTTCCATATAAGGAGTACAAAGCAAGTTGGAAAGTCAACAAACGGAAGGA 395			
Qy	361	TTTAAACAAGGATTTGGGAAATAGAAATAACCCAGGAGTAAAGTTACTGCTACCG 420			
Db	396	TTTAAACAAGGATTTGGGAAATAGAAATAACCCAGGAGTAAAGTTACTGCTACCG 455			
Qy	421	GCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGAGCAAGTGGAAATATCGCAGATGCA 480			

Db 456 GCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGAGAAAGTGGAATACTGCAGATGCA 515
 Qy 481 AGCAGTGAAGGAAGAGGTGATAGATAGAGAGAGATGAAAGGCAAAAGAAAGAAATGAA 540
 Db 516 AGCAGTGAAGGAAGAGGTGATAGATAGAGAGAGATGAAAGGCAAAAGAAAGAAATGAA 575
 Qy 541 AAAGCAGCTCAAAACGGAAAAAGCTATATATCTTCAAGAAATCCTCTAAACAGTCCCGG 600
 Db 576 AAAGCAGCTCAAAACGGAAAAA-TCATATATCTTCAAGAAATCCTCTAAACAGTCCCGG 634
 Qy 601 AAATCTCCAGG-AGATGAAGATGACAAAGACTGCAAGAGAGAGAGAAACAAACAGCAGCTC 659
 Db 635 AAATCTCCAGGAGATGAAGATGACAAAGACTGCCNAGAGAAAGAAACAAACAGCAGCTC 694
 Qy 660 TGAGGTGGAGATG-CGGGCAACGACACAAAGAAACAACTTC 701
 Db 695 TGAAGTGGANATCGCGGCAACGACACACNAGAAACACACTTC 737

RESULT 13
 AB029493
 LOCUS
 DEFINITION Mus musculus mRNA for hepatoma-derived growth factor-related protein HRP-3, complete cds.
 ACCESSION AB029493
 VERSION AB029493.1 GI:6855469
 KEYWORDS hepatoma-derived growth factor-related protein HRP-3.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Ikegane,K., Yamamoto,M., Kishima,Y., Enomoto,H., Yoshida,K., Shemura,M., Kishimoto,T. and Nakamura,H.
 TITLE A new member of a hepatoma-derived growth factor gene family can translocate to the nucleus
 JOURNAL Biochem. Biophys. Res. Commun. 266 (1), 81-87 (1999)
 MEDLINE 20050055
 PUBMED 10581169
 REFERENCE 2 (bases 1 to 612)
 AUTHORS Ikegane,K. and Nakamura,H.
 TITLE Direct Submission
 JOURNAL Submitted (01-JUL-1999) Kazuhiro Ikegane, Osaka University Medical School, Department of Molecular Medicine; 2-2 Yamada-oka, Suita, Osaka 565-0871, Japan (E-mail:nakamura@imed3.med.osaka-u.ac.jp, Tel:81-6-6879-3837(ex.3837), Fax:81-6-6879-3839)

FEATURES
 source Location/Qualifiers
 1..612
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /tissue_type="testis"
 1..612
 /gene="HRP-3"
 1..609
 /gene="HRP-3"
 /codon_start=1
 /product="hepatoma-derived growth factor-related protein HRP-3"
 /protein_id="BAA90478.1"
 /db_xref="GI:6855470"
 /translation="MARPRREYKAGDLVFAKMGYPHPARIDELPEGAVKPPANKYPIFFGTHETAFGLPKDLFPYKEVKQKFGSKNKRKGFNEGLWEINNPVGKFTGYGTIQQSSSTEGEGNTADASSEEDGRVEDGKGRKQKGGSKRKYSTKSKSQSRKSPGEDKCKEENKSSSGGNDTRNTADLQKTSBT"

BASE COUNT
 ORIGIN 223 a 123 c 161 g 105 t

Query Match 53.5%; Score 547.8; DB 10; Length 612;
 Best Local Similarity 94.3%; Pred. No. 1.5e-123;
 Matches 580; Conservative 0; Mismatches 32; Indels 3; Gaps 1;

Qy 121 ATGCGCGTCCGCGCCCGCGAGTACAAAGCGGCGGCGACCTGGTCTTCGCCAAGATGAAG 180

Db 1 ATGCGCGTCCGCGCCCGCGAGTACAAAGCGGAGACCTGGTCTTCGCCAAGATGAAG 60
 Qy 181 GGCTACCCGACATGGCGCGCCCGGATGATGAATCTCCAGAGGCGCTGTGAAGCTTCCA 240
 Db 61 GGCTACCCGACATGGCGCGCCCGGATGATGAATCTCCAGAGGAGCTGTGAAGCTTCCA 120
 Qy 241 GCAAAACAAGTATCTATCTCTCTTTTGGCACCACCTGAAACTGCATTTCTAGGTCCCAAA 300
 Db 121 GCAACAAGTATCTATCTCTCTTTTGGTACCCATGAACTGCATTTCTAGGTCTTAAA 180
 Qy 301 GACCTTTTCCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAACAAACGGAAGGA 360
 Db 181 GACCTTTTCCATATAGGAAATACAAAGCAAGTTTGGAAAGTCAACAAACGGAAGGA 240
 Qy 361 TTTTAAAGAGGATTTGGGAAATAGAAATAAACCAGGAGTAAAGTTTACTGGCTACCAG 420
 Db 241 TTTTAAAGAGGATTTAGGAAATTTGAAATAAATCCAGGAGTGAATTTTACTGGGTACCAG 300
 Qy 421 GCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGAGAGAGGTGGAATACTGCAGATGCA 480
 Db 301 ACATTCAGCAACAGAGCTCTTCAGAAACTGAGGAGAGAGGTGGAATACTGCAGATGCA 360
 Qy 481 AGCAGTGAGGAAGAAGGTGATAGAGTAGAAGAGATGGAAGGCAAAAGAAAGAAATGAA 540
 Db 361 AGCAGTGAGGAAGAAGGTGATAGAGT--AGAAGATGGAAGGCAAAAGAAAGAAATGAA 417
 Qy 541 AAAGCAGGCTCAAAACGGAAAAAGTATATATCTTCAAGAAATCTCTTAAACAGTCCCGG 600
 Db 418 AAAGCAGGCTCAAAACGGAAAAAGTCTTACACGCTCAAGAAAGTCTTCTTAAACAGTCCCGG 477
 Qy 601 AAATCTCCAGGAGATGAAGATGACAAAGACTGCAAGAGAGAGGAAACAAACAGCAGCTCT 660
 Db 478 AAATCTCCAGGAGATGAAGATGATTAAGACTGCAAGAGAGAGGAAACAAACAGCAGCTCT 537
 Qy 661 GAGGCTGGAGATCGCGGCAACGACACAAAGAAACACAACTTTCAGACTTTCAGAGAAACCCAGT 720
 Db 538 GAGGCGGAGAGCTGGCAATGACACAGAAACACGACTGCAGACTTTCAGAGAAACCCAGT 597
 Qy 721 GAAGGAGACCTAACTA 735
 Db 598 GAAGGAGACCTAACTA 612

RESULT 14
 IRO783648
 LOCUS Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 783648.
 DEFINITION Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 783648.
 ACCESSION AL109689
 VERSION AL109689.1 GI:5689859
 KEYWORDS FLI_CDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 Auffray,C., Ansoorge,W., Ballabio,A., Estivill,X., Gibson,K., Lehrach,H., Poustka,A. and Dundeberg,U.
 TITLE The European IMAGE consortium for integrated Molecular analysis of human gene transcripts
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1247)
 AUTHORS Carim,L., Estivill,X., Sumoy,L. and Escarceller,M.
 TITLE Direct Submission
 JOURNAL Submitted (31-JUL-1999) Departament de Genetica Molecular, Institut de Recerca Oncologica (IRO), Hospital Duran i Reynals, Autovia de Castelldefels Km 2,7 L'Hospitalet de Llobregat, 08907 Barcelona, Catalunya, SPAIN. Tel: ++34-93-260-7775 Fax: ++34-93-260-7776 WWW site: http://www.iro.es e-mail enquires: lsumoy@iro.es, mescarceller@iro.es

COMMENT
 EURO-IMAGE Consortium Contact: Auffray C
 CNRS UPR 420 - Genetique Molculaire et Biologie du Developement
 IFR 1221 - Rue Guy Moquet 19, Batiment G - BP 8

94801 Villejuif Cedex, FRANCE
Tel: ++33-1-49 58 34 98
Fax: ++33-1-49 58 35 09
e-mail: aufrey@infobiogen.fr

This clone is available royalty-free through IMAGE Consortium Distributors.

IMPORTANT: This sequence represents the full insert of this IMAGE CDNA clone. No attempt has been made to verify whether this corresponds to the full-length of the original mRNA from which it was derived.

FEATURES

Location/Qualifiers

1..1247
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosomes="15"
/map="15q24-q25"
/clone="IMAGE cDNA clone 783648"
/clone_lib="Soares_total_fetus_Nb2HF8_9w"
-1..335
/codon_start=3
/product="CDI-142 hypothetical protein"
/protein_id="CB52021.1"
/db_xref="GI:5689860"
/translation="NLOEMKMTKTKAKRKTAAALRVEMRATTQETLOTCRKPVKGN
YNECCILRETTRELYVWLSNLILGDMQHIVLVDRTPVCMYIIHPLCCVSGGKD
ILAFKSY"

CDS

BASE COUNT 427 a 197 c 209 g 414 t
ORIGIN

Query Match 41.5%; Score 425; DB 9; Length 1247;

Best Local Similarity 100.0%; Pred. No. 2.5e-93; Mismatches 0; Indels 0; Gaps 0;

Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 600 GAAATCTCCAGGAGATGAAGATGACAAAGACTGCAAAAGAGAGGAAACAAAGCAGCTC 659

Db 1 GAAATCTCCAGGAGATGAAGATGACAAAGACTGCAAAAGAGAGGAAACAAAGCAGCTC 60

QY 660 TGAGGTGAGATGCGGGCAACGACACAAAGAAACAACTTCAGACTTCGAGAAACCCAG 719

Db 61 TGAGGTGAGATGCGGGCAACGACACAAAGAAACAACTTCAGACTTCGAGAAACCCAG 120

QY 720 TGAAGGGACCTAACTACCATATGATGCTGCATATTAGAGAAACCAAGAGGTTAT 779

Db 121 TGAAGGGACCTAACTACCATATGATGCTGCATATTAGAGAAACCAAGAGGTTAT 180

QY 780 ATGTTGGTGTCTAATATCTTGGATTTGATATGAACCAACATAGTCCTTGTGTCA 839

Db 181 ATGTTGGTGTCTAATATCTTGGATTTGATATGAACCAACATAGTCCTTGTGTCA 240

QY 840 TTGACAGAACCCAGTTGTATGATATTCATATTCCTCTCTGTTGTGTTTCGGGG 899

Db 241 TTGACAGAACCCAGTTGTATGATATTCATATTCCTCTCTGTTGTGTTTCGGGG 300

QY 900 GAAAGACATTTAGCTTTTAAAGTTACTGATTTAATTTCAATTTATTTGTTGCA 959

Db 301 GAAAGACATTTAGCTTTTAAAGTTACTGATTTAATTTCAATTTATTTGTTGCA 360

QY 960 TGAAGTGCCTTAAACCACTAAGGATTAATCAAGATTTTTCGACAGATTATACATGTCTA 1019

Db 361 TGAAGTGCCTTAAACCACTAAGGATTAATCAAGATTTTTCGACAGATTATACATGTCTA 420

QY 1020 GGATC 1024

Db 421 GGATC 425

RESULT 15

AL590240/c

LOCUS

DEFINITION

AL590240

ACCESSION

Human DNA sequence from clone RP11-26613 on chromosome X, complete

sequence.

AL590240 AC025380

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL590240.5 GI:13872461

HTG

Homo sapiens (human)

Homo sapiens

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 130005)

Clark, S.

Direct Submission

Submitted (25-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

On or before May 15, 2001 this sequence version replaced

gi:8705130, gi:13785125.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. The following

abbreviations are used to associate primary accession numbers given

in the feature table with their source databases: Em, EMBL; Sw,

SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP

database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence

was generated from part of bacterial clone contigs of human

chromosome X, constructed by the Sanger Centre Chromosome X Mapping

Group. Further information can be found at

http://www.sanger.ac.uk/HGP/ChrX

RP11-26613 is from the library RP11-11.1 constructed by the group

of Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm

VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone

RP11-26613 It may be shorter because we sequence overlapping

sections only once, except for a 100 base overlap.

The true left end of clone RP11-382F24 is at 129906 in this

sequence. The true right end of clone RP1-296K21 is at 100 in this

sequence.

Location/Qualifiers

1..130005

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="X"

/clone_lib="RP11-26613"

/clone_lib="RP11-11.1"

1..3875

/note="L1P repeat: matches 4..3892 of consensus"

4031..4472

/note="L1PA10 repeat: matches 5718..6165 of consensus"

4477..4658

/note="L1ME3A repeat: matches 5967..6157 of consensus"

4869..5170

/note="AluSc repeat: matches 1..306 of consensus"

6010..6307

/note="AluSx repeat: matches 1..299 of consensus"

9580..9670

/note="WIR repeat: matches 127..214 of consensus"

10019..10387

/note="WL1B repeat: matches 41..390 of consensus"

10424..11726

/note="L1MA7 repeat: matches 4980..6284 of consensus"

12183..12372

/note="L1PA15 repeat: matches 5963..6153 of consensus"

12387..17012

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2004, 06:07:00 ; Search time 328 Seconds
(without alignments)
8427.514 Million cell updates/sec

Title: US-09-787-328B-3

Perfect score: 1024

Sequence: 1 accgcgtccgcgcgcgcgtt.....acttatacatgtctaggatc 1024

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :				N_Geneseq_19Jun03.*
1:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1980.DAT.*			
2:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1981.DAT.*			
3:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1982.DAT.*			
4:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1983.DAT.*			
5:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1984.DAT.*			
6:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1985.DAT.*			
7:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1986.DAT.*			
8:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1987.DAT.*			
9:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1988.DAT.*			
10:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1989.DAT.*			
11:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1990.DAT.*			
12:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1991.DAT.*			
13:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1992.DAT.*			
14:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1993.DAT.*			
15:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1994.DAT.*			
16:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1995.DAT.*			
17:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1996.DAT.*			
18:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1997.DAT.*			
19:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1998.DAT.*			
20:	/SIDSI/gcgdata/geneq/geneq-emb1/NA1999.DAT.*			
21:	/SIDSI/gcgdata/geneq/geneq-emb1/NA2000.DAT.*			
22:	/SIDSI/gcgdata/geneq/geneq-emb1/NA2001A.DAT.*			
23:	/SIDSI/gcgdata/geneq/geneq-emb1/NA2001B.DAT.*			
24:	/SIDSI/gcgdata/geneq/geneq-emb1/NA2002.DAT.*			
25:	/SIDSI/gcgdata/geneq/geneq-emb1/NA2003.DAT.*			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1024	100.0	1024	21	AA13161 Human type II hepa
2	1012	98.8	1973	22	AAH14117 Human cDNA sequenc
3	1000	97.7	1898	22	AAF76854 Human secreted pro
4	985.6	96.2	1919	21	AACT7751 Human cancer assoc
5	980	95.7	3805	20	AAZ00048 HGF2 gene. Homo
6	969	94.6	4487	24	ABST70439 Human bone remodel
7	828.4	80.9	1664	24	ABT06906 Human ovarian can
8	828.4	80.9	1664	24	ABT07105 Human ovarian can

9	828.4	80.9	1664	25	ABX72784 Human ovarian carc
10	828.4	80.9	1664	25	ABX72983 Human ovarian carc
11	778	76.0	2920	24	ABL35012 Rat cDNA isolated
12	623.6	60.9	737	22	AAH05630 Human cDNA clone (
13	609	59.5	609	21	AAAG5004 Human hepatoma der
14	555.4	54.2	870	23	AAAG79520 DNA encoding novel
15	471	46.0	645	21	AAZ80234 Human colon cancer
16	352.6	34.4	596	23	AAAG79519 Human immune/haema
17	298	29.1	5744	22	AAK86342 Human secreted pro
18	263.6	25.7	334	21	AAAC31156 Bovine EST associa
19	248.6	24.3	398	25	ABX43447 Bovine EST associa
20	239.6	23.4	288	25	ABX5860 Human HDGF nucleo
21	193.4	18.9	1386	22	AAAF5292 Human HDGF nucleo
22	193.4	18.9	2376	16	AAQ79902 Human hepatoma der
23	193.4	18.9	2376	21	AAAG40118 Human HDGF cDNA.
24	193.4	18.9	2376	22	ABAG3087 Hepatoma-derived g
25	189	18.5	723	16	AAQ79903 Human hepatoma der
26	189	18.5	2133	21	AAAF18049 Lung cancer associ
27	186.6	18.4	869	19	AAAV39154 Lung growth factor
28	186.6	18.2	732	21	AAAG40112 Human HDGF-Ib cDNA
29	186.6	18.2	834	21	AAAG40117 Human HDGFIBN cDNA
30	183.8	17.9	714	18	AAT88420 Mouse hepatoma der
31	183.8	17.9	1563	18	AAV03867 Mouse hepatoma der
32	174.6	17.1	3117	24	AAAS18198 Human nuclear ancl
33	174.6	17.1	3117	24	AAAS20482 Human lens epithel
34	174.6	17.1	3117	24	AAAS20483 Human nuclear tran
35	173	16.9	1001	21	AAAG27153 5' region of human
36	173	16.9	1763	21	AAAG27151 Human p52 cDNA. H
37	173	16.9	2347	21	AAAG27150 Human p75 cDNA. H
38	173	16.9	3360	20	AAAG7285 Human lens epithel
39	172.6	16.9	800	24	ABST7354 Frog embryonic gen
40	172.6	16.9	800	24	ABST7359 Frog embryonic gen
41	171	16.7	2289	24	ABN86075 Human erythrocyte
42	163.4	16.0	2108	19	AAV01731 Mouse liver cancer
43	159.8	15.6	983	22	ABAG9360 Human secreted pro
44	159.8	15.6	1297	22	AAAF76856 Human NOV9 coding
45	159.8	15.6	2031	22	ABAG1999 Human NOV9 coding

ALIGNMENTS

RESULT 1	
AA13161	
ID	AA13161 standard; cDNA; 1024 BP.
AC	AA13161;
XX	
XX	20-JUL-2000 (first entry)
XX	
XX	Human type II hepatoma derived growth factor (HDGF2) nucleotide sequence.
DE	
XX	Hepatoma-derived growth factor 2; HDGF2; cytostatic; drug;
KW	treatment; hepatoma; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200017351-A1.
XX	
PD	30-MAR-2000.
XX	
PF	06-SEP-1999; 99WO-CN00139.
XX	
PR	22-SEP-1998; 98CN-0119758.
XX	
PA	(YULL/) YU L.
XX	
PI	Yu L, Zhang H, Fu Q, Zhao Y, Tu Q;
XX	
DR	WPI; 2000-283579/24.
DR	P-PSDB; AAY88374.
XX	
PT	Type II human-derived growth factor with homology to its type I

PT version, useful e.g. in study and development of drugs for hepatoma -
XX Claim 1; Page 14; 27pp; Chinese.
XX This sequence represents the cDNA sequence encoding the human
CC hepatoma-derived growth factor (HDGF2) protein. The HDGF2 sequence shows
CC considerable homology to the type I version. The protein has cytostatic
CC activity. The HDGF2 protein, derivatives and polynucleotides are useful
CC e.g. in the study and development of drugs for the treatment of hepatoma.
XX
SQ Sequence 1024 BP; 323 A; 215 C; 262 G; 224 T; 0 other;
Query Match 100.0%; Score 1024; DB 21; Length 1024;
Best Local Similarity 100.0%; Pred. No. 1.3e-263;
Matches 1024; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCGCTCGTCCGCGCTTGAAGCGCGCGGAGCGCGCAATTCGTCGCGCGCGGG 60
DB 1 ACCGCTCGTCCGCGCTTGAAGCGCGCGGAGCGCGCAATTCGTCGCGCGCGGG 60
QY 61 GGGGCGGCTCCCGGCACTTCGCGCGCAACCAAGCACTACCAAGGAGGGAGCGCTGG 120
DB 61 GGGGCGGCTCCCGGCACTTCGCGCGCAACCAAGCACTACCAAGGAGGGAGCGCTGG 120
QY 121 ATGGCGCGTCCGCGCGCGCGAGTACAAGCGCGGCGACCTGCTCTTCGCCAAGATGAAG 180
DB 121 ATGGCGCGTCCGCGCGCGCGAGTACAAGCGCGGCGACCTGCTCTTCGCCAAGATGAAG 180
QY 181 GGTACCCGCACTGGCGCGCGCGCGATTTGATGAATCCCAAGGCGCGCTGTGAAGCTCCA 240
DB 181 GGTACCCGCACTGGCGCGCGCGCGATTTGATGAATCCCAAGGCGCGCTGTGAAGCTCCA 240
QY 241 GCAACACAGTATCTTCTCTCTTTTGGCAGCCATGAATCTTCTAGGTCCTCCAA 300
DB 241 GCAACACAGTATCTTCTCTCTTTTGGCAGCCATGAATCTTCTAGGTCCTCCAA 300
QY 301 GACCTTTTCCATATAAGAGTACAAAGCAAGATTTGAAAGTCAAAAGCAAGAGGA 360
DB 301 GACCTTTTCCATATAAGAGTACAAAGCAAGATTTGAAAGTCAAAAGCAAGAGGA 360
QY 361 TTTAAGCAAGAGTGTGGAAATAGAAATATACCCAGGAGTAAAGTTTACTGGCTACCAG 420
DB 361 TTTAAGCAAGAGTGTGGAAATAGAAATATACCCAGGAGTAAAGTTTACTGGCTACCAG 420
QY 421 GCAATTCAGCAACAGAGCTTTCAGAACTGAGGAGAGGTGGAATTAATCTCAGATGCA 480
DB 421 GCAATTCAGCAACAGAGCTTTCAGAACTGAGGAGAGGTGGAATTAATCTCAGATGCA 480
QY 481 AGCAGTGAGGAAGAGTGTATAGATAGAGAGAGTGGAAAGGCAAAAGAAAGATGAA 540
DB 481 AGCAGTGAGGAAGAGTGTATAGATAGAGAGAGTGGAAAGGCAAAAGAAAGATGAA 540
QY 541 AAAGCAGGCTCAAAACGGAAGAGTATATCTTCAAGAAATCTCTTAAACAGTCCCGG 600
DB 541 AAAGCAGGCTCAAAACGGAAGAGTATATCTTCAAGAAATCTCTTAAACAGTCCCGG 600
QY 601 AAATCTCAGGAGATGAAGTACAAAGCTCAAGAGAGAGAGAGAGAGAGAGTCT 660
DB 601 AAATCTCAGGAGATGAAGTACAAAGCTCAAGAGAGAGAGAGAGAGAGTCT 660
QY 661 GAGGTGGAGATGCGGGCAACACACAAAGCAACACTTCAGACTTCGAGAAACCAAGT 720
DB 661 GAGGTGGAGATGCGGGCAACACACAAAGCAACACTTCAGACTTCGAGAAACCAAGT 720
QY 721 GAAGGAGACCTAATCAATGAATGCTGCATATTAAGAGAAACCAAGAGAGTATA 780
DB 721 GAAGGAGACCTAATCAATGAATGCTGCATATTAAGAGAAACCAAGAGAGTATA 780
QY 781 TGTGTGGTGTCTAATATCTTGGATTTGATATGAACCAACATAGTCTTGTGTCTAT 840
DB 781 TGTGTGGTGTCTAATATCTTGGATTTGATATGAACCAACATAGTCTTGTGTCTAT 840
QY 841 TGACAGAACCCAGTTTGTATGATATATTCATATTCCTCTCTCTGTTGTTTCGGGGG 900

DB 841 TGACAGAACCCAGTTTGTATGATATATTCATATTCCTCTCTGTTGTTTCGGGGG 900
QY 901 AAAAGACATTTAGCCCTTTTAAAGTTACTGATTTAAATTTTCAATTTTGGTTGCAT 960
DB 901 AAAAGACATTTAGCCCTTTTAAAGTTACTGATTTAAATTTTCAATTTTGGTTGCAT 960
QY 961 GAAGTTGGCTTTAACCACCTAAGGATTTATCAAGATTTTGGCAGACTTTATACATCTCTAG 1020
DB 961 GAAGTTGGCTTTAACCACCTAAGGATTTATCAAGATTTTGGCAGACTTTATACATCTCTAG 1020
QY 1021 GATC 1024
DB 1021 GATC 1024
RESULT 2
AAH14117
ID AAH14117 standard; cDNA; 1973 BP.
XX AC AAH14117;
XX 26-JUN-2001 (first entry)
XX Human cDNA sequence SEQ ID NO:11306.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-0116126.
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs -
XX Claim 8; SEQ ID 11306; 2537pp + CD ROM; English.
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dr primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length

Best Local Similarity 99.8%; Pred. No. 4.6e-257;		Matches 1022; Conservative 0; Mismatches 0; Indels 2; Gaps 2;	
QY	2	CGCTCGTCCGCGCGCTTTGAGCCCGCGGGAGCGC-GGCAATTCGTGCGCGCGCGG 60	
Db	30	CGCTCGTCCGCGCGCTTTGAGCCCGCGGGAGCGCGCAATTCGTGCGCGCGCGG 89	
QY	61	GGGCGGCGCTCCCGGCATCTTCGCGCGACCAAGGACTACCAAGAAAGGGAGCGGCTGG 120	
Db	90	GGGCGGCGCTCCCGGCATCTTCGCGCGACCAAGGACTACCAAGAAAGGGAGCGGCTGG 149	
QY	121	ATGCGCGCTCCGCGCGCCCGCGAGTACAAAGCGCGGACCTGCTCTTCGCCAAGATGAAG 180	
Db	150	ATGCGCGCTCCGCGCGCCCGCGAGTACAAAGCGCGGACCTGCTCTTCGCCAAGATGAAG 209	
QY	181	GGCTACCGGACTGCGCGCGCGCGGATGATGAATCCAGAGGCGCTGGAAGCTCCA 240	
Db	210	GGCTACCGGACTGCGCGCGCGCGGATGATGAATCCAGAGGCGCTGGAAGCTCCA 269	
QY	241	GCAACAAGTATCTATCTCTTTTGGCACCCTGAACTGCAATTCCTAGTCCCAA 300	
Db	270	GCAACAAGTATCTATCTCTTTTGGCACCCTGAACTGCAATTCCTAGTCCCAA 329	
QY	301	GACCTTTTCCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAGGA 360	
Db	330	GACCTTTTCCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAGGA 389	
QY	361	TTTAACGAAGATTGGGGAATAGAAATTAACCCAGGAGTAAAGTTTACTGCTACCA 420	
Db	390	TTTAACGAAGATTGGGGAATAGAAATTAACCCAGGAGTAAAGTTTACTGCTACCA 449	
QY	421	GCAATTCAGCAACAGAGCTCTTCAGAACTTGAGGAGAGAGTGGAAATCTCAGATGA 480	
Db	450	GCAATTCAGCAACAGAGCTCTTCAGAACTTGAGGAGAGAGTGGAAATCTCAGATGA 509	
QY	481	AGCAGTGAGGAAAGGTTGATAGTAGAGTAGAAGATGGAAGGCAAGAAAGAAATGAA 540	
Db	510	AGCAGTGAGGAAAGGTTGATAGTAGAGTAGAAGATGGAAGGCAAGAAAGAAATGAA 569	
QY	541	AAAGCAGGCTCAAAACGGAAGATCATATCTTCAAGAAATCTCTAAACAGTCCCG 600	
Db	570	AAAGCAGGCTCAAAACGGAAGATCATATCTTCAAGAAATCTCTAAACAGTCCCG 629	
QY	601	AAATCTCCAGGAGTCAAGATGACAAAGACTGCAAGAGAGAGAAACAAAGAGCTCT 660	
Db	630	AAATCTCCAGGAGTCAAGATGACAAAGACTGCAAGAGAGAGAAACAAAGAGCTCT 689	
QY	661	GAGGTGGAGATGCGGGCAACGACACAGAAACACAACTTCAGACTTCGAGAAACCAAGT 720	
Db	690	GAGGTGGAGATGCGGGCAACGACACAGAAACACAACTTCAGACTTCGAGAAACCAAGT 749	
QY	721	GAAGGACCTAATCACTACCAATAGATGCTGCATATTAAGAGAAACCAAGAGGTTATA 780	
Db	750	GAAGGACCTAATCACTACCAATAGATGCTGCATATTAAGAGAAACCAAGAGGTTATA 809	
QY	781	TGTTTGTTGCTCTAATATCTTGGATTGATATGACCAACACATAGTCTCTTGTTGTCAT 840	
Db	810	TGTTTGTTGCTCTAATATCTTGGATTGATATGACCAACACATAGTCTCTTGTTGTCAT 869	
QY	841	TGACAGAACCCAGTTTGTATGTATATTAATTCATATTCCTCTCTGTGTGTTTCGGGGG 900	
Db	870	TGACAGAACCCAGTTTGTATGTATATTAATTCATATTCCTCTCTGTGTGTTTCGGGGG 929	
QY	901	AAAAGACATTTAGCTTTTAAAGTACTGATTAATTAATTCATGTTATTTGGTTCAT 960	
Db	930	AAAAGACATTTAGCTTTTAAAGTACTGATTAATTAATTCATGTTATTTGGTTCAT 989	
QY	961	GAAGTTGCCCTTAACCACTAAGGATTATCAAGATTTTTCGCGAGACTTATACATGCTAG 1020	
Db	990	GAAGTTGCCCTT-ACCACCTAAGGATTATCAAGATTTTTCGCGAGACTTATACATGCTAG 1048	
QY	1021	GATC 1024	

Db	1049	GATC 1052	
RESULT 4			
AAC77751	ID AAC77751 standard; cDNA; 1919 BP.		
XX	AC	AAC77751;	
XX	08-FEB-2001	(first entry)	
XX	Human cancer associated gene sequence SEQ ID NO:145.		
DE	Human; cancer associated gene; cancer antigen; detection; cancer;		
KW	diagnosis; cytosolic; proliferative; cancer; antigen; immunomodulator;		
KW	antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;		
KW	antifungal; antiparasitic; antitubercular; antibacterial; cardiac;		
KW	dermatological; neuroprotective; thrombolytic; coagulant; neoplastic;		
KW	vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;		
KW	immune disorder; haematopoietic cell disorder; autoimmune disorder;		
KW	allergic reaction; graft versus host disease; organ rejection;		
KW	haemostatic; thrombolytic; cardiovascular disorder; infection;		
KW	neurological disease; drug screening; ss.		
OS	Homo sapiens.		
XX	WO200055350-A1.		
XX	21-SEP-2000.		
XX	08-MAR-2000; 2000WO-US05882.		
XX	12-MAR-1999; 99US-0124270.		
XX	(HUMA-) HUMAN GENOME SCI INC.		
XX	Rosen CA, Ruben SM;		
XX	WPI: 2000-587533/55.		
XX	P-PSDB; AAB43542.		
XX	Novel isolated nucleic acids comprising sequences encoding peptides		
XX	useful for treating or diagnosing e.g. cancer -		
XX	Claim 1; Page 724-725; 2352pp; English.		
XX	AAC77607 to AAC78448 encode the human cancer associated proteins given		
XX	in AAB43398 to AAB44239. The proteins can have activities based on the		
XX	tissues and cells the genes are expressed in. Example of activities		
XX	include: cytostatic; proliferative; vulnary; immunomodulator;		
XX	antidiabetic; antiasthmatic; antirheumatic; antiarthritic;		
XX	antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;		
XX	dermatological; neuroprotective; cardiac; thrombolytic; coagulant;		
XX	neoplastic; vasotropic; antipsoriatic and antiangiogenic. The		
XX	polynucleotides and polypeptides can be used for preventing, treating or		
XX	ameliorating medical conditions and diagnosing pathological conditions.		
XX	Polynucleotides, polypeptides, antibodies, agonists and antagonists from		
XX	the present invention may be used to treat immune disorders by activating		
XX	or inhibiting the proliferation, differentiation or mobilisation of		
XX	immune cells, to treat disorders of haematopoietic cells, autoimmune		
XX	disorders, allergic reactions, graft versus host disease and organ		
XX	rejection, modulate haemostatic or thrombolytic activity, modulate		
XX	inflammation, cancer, cardiovascular disorders, neurological disease and		
XX	bacterial or viral infections. The peptides, nucleotides, antibodies, to		
XX	agonists and antagonists may be also used in drug screens. AAC78449 to		
XX	AAC78457 and AAB44240 represent sequences used in the exemplification of		
XX	the present invention.		
XX	Sequence 1919 BP; 613 A; 352 C; 409 G; 539 T; 6 other;		
SQ	Query Match 96.2%; Score 985.6; DB 21; Length 1919;		
	Best Local Similarity 99.4%; Pred. No. 3.3e-253;		
	Matches 1018; Conservative 2; Mismatches 1; Indels 3; Gaps 3;		

QY 2 CCCTCTGTCGCGCGGCTTGAAGCCCGCGGGAGCGC-GCGCAATTGTCGCGCCCGCGG 60
Db |||||
QY 15 CCCTCTGTCGCGCGGCTTGAAGCCCGCGGGAGCGGCGCAATTGTCGCGCCCGCGG 74
Db |||||
QY 61 GGGCGGCGCTCCCGGCATCTTCGGGCGGACCAAGGACTACAGGAAGGGAGCGGCTGG 120
Db |||||
QY 75 GGGCGGCGCTCCCGGCATCTTCGGGCGGACCAAGGACTACAGGAAGGGAGCGGCTGG 134
Db |||||
QY 121 ATGGCGCGTCCGGCGCCCGCGAGTACAAAGCGGCGGACCTGCTCTTCGCAAGATGAAG 180
Db |||||
QY 135 ATGGCGGCTCG-GGCCCGSAGTACAAAGCGGCGGAGCTGCTCTTCGCAAGATGAAG 193
Db |||||
QY 181 GGCTACCGCACTGGCGCGCCCGGATGATGAATCCAGAGGCGCTGTGAAGCCCTCCA 240
Db |||||
QY 194 GGCTACCGCACTGGCGCGCCCGGATGATGAATCCAGAGGCGCTGTGAAGCCCTCCA 252
Db |||||
QY 241 GCAACCAAGTATCTATCTCTCTTTTGGGACCCATGAAGCTGATTTCTAGTCCCAA 300
Db |||||
QY 253 GCAACCAAGTATCTATCTCTTTTGGGACCCATGAAGCTGATTTCTAGTCCCAA 312
Db |||||
QY 301 GACCTTTTCCATATAGAGTACAAAGACAAAGTTTGGAAAGTCAACCAACGGAAGGA 360
Db |||||
QY 313 GACCTTTTCCATATAGAGTACAAAGACAAAGTTTGGAAAGTCAACCAACGGAAGGA 372
Db |||||
QY 361 TTTAACGAAGGATTTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACCAG 420
Db |||||
QY 373 TTTAACGAAGGATTTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACCAG 432
Db |||||
QY 421 GCAATTTCAGCAAGAGCTCTTCAGAAACTGAGGAGAGGAGTGAATATCTGAGATGCA 480
Db |||||
QY 433 GCAATTTCAGCAAGAGCTCTTCAGAAACTGAGGAGAGGAGTGAATATCTGAGATGCA 492
Db |||||
QY 481 AGCAGTGAGCAAGAGTGTAGAGTAGAGAGATGGAAGGCAAGGAGGAGGAGATGA 540
Db |||||
QY 493 AGCAGTGAGCAAGAGTGTAGAGTAGAGAGATGGAAGGCAAGGAGGAGGAGATGA 552
Db |||||
QY 541 AAACGAGGCTCAAAACGGAAGAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCCGG 600
Db |||||
QY 553 AAACGAGGCTCAAAACGGAAGAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCCGG 612
Db |||||
QY 601 AATCTCCAGGAGTGAAGTGAACAAAGCTGCAAGAGAGGAGGAGGAGGAGGAGTCT 660
Db |||||
QY 613 AATCTCCAGGAGTGAAGTGAACAAAGCTGCAAGAGAGGAGGAGGAGGAGGAGTCT 672
Db |||||
QY 661 GAGGCTGGAGTCCGCGCAACGACACAAAGAAACACAACTTCAGACTTCGAGGAAACCACT 720
Db |||||
QY 673 GAGGCTGGAGTCCGCGCAACGACACAAAGAAACACAACTTCAGACTTCGAGGAAACCACT 732
Db |||||
QY 721 GAAGGACCTTAACCATTAATGAATGCTGCATATTAAAGAAACCAAGAAAGGTTATA 780
Db |||||
QY 733 GAAGGACCTTAACCATTAATGAATGCTGCATATTAAAGAAACCAAGAAAGGTTATA 792
Db |||||
QY 781 TGTTGGTGTCTTAATATCTTGGATTTGATGAACCAACACATAGTCTTGTGTCAT 840
Db |||||
QY 793 TGTTGGTGTCTTAATATCTTGGATTTGATGAACCAACACATAGTCTTGTGTCAT 852
Db |||||
QY 841 TGACAGAACCCAGTTGTATGTACATATTCTCTCTCTGTGTGTGTTTCGGGGGG 900
Db |||||
QY 853 TGACAGAACCCAGTTGTATGTACATATTCTCTCTCTGTGTGTGTTTCGGGGGG 912
Db |||||
QY 901 AAAAGACATTTTAGCCCTTTTAAAAAGTTACTGATTTAATTTTCATGTTTGGTTGCAT 960
Db |||||
QY 913 AAAAGACATTTTAGCCCTTTTAAAAAGTTACTGATTTAATTTTCATGTTTGGTTGCAT 972
Db |||||
QY 961 GAAGTTGCCCTTAAACCAATGAAGATTATCAAGATTTTTCGCGAGACTTATACATGCTAG 1020
Db |||||
QY 973 GAAGTTGCCCTTAAACCAATGAAGATTATCAAGATTTTTCGCGAGACTTATACATGCTAG 1032
Db |||||
QY 1021 GATC 1024
Db |||||
QY 1033 GATC 1036

RESULT 5
AAZ00048
ID AAZ00048 standard; DNA; 3805 BP.
XX
AC AAZ00048;
XX
DT 19-OCT-1999 (first entry)
XX
DE HGFH2 gene.
XX
XX
KW HGFH2, human growth factor homologue 2; cell proliferation; cancer;
human hepatoma derived growth factor; HDGF; hepatitis; anaemia; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 134..733
FT /*tag= a
FT /product= "HGFH2
FT /note= "Human growth factor homologue 2"
XX
PN WO938976-R2.
XX
PD 05-AUG-1999.
XX
PP 12-JAN-1999; 99WO-US000654.
XX
PR 29-JAN-1998; 98US-0015412.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Baughn M, Corley NC, Guegler KJ, Hillman JL, Lal P;
PI Tang Yt;
XX
DR WPI; 1999-469331/39.
DR P-PSDB; AAY28502.
XX
PT New human growth factor homologs and their use in the diagnosis,
treatment and prevention of cell proliferative and developmental
disorders
PT
PS
XX
Claim 7; Page 77-78; 88pp; English.
XX
CC This sequence is the HGFH2 Human Growth Factor Homologue 2 gene. The
HGFH2 protein AAY28502 has structural and chemical homology with the
human hepatoma derived growth factor (HDGF). HGFH2s are expressed in
cancerous tissues and appear to play a role in cell proliferative and
developmental disorders. A purified antagonist of HGFH2 can be
administered to a subject as a method of treating or preventing a cell
proliferative disorder e.g. atherosclerosis, hepatitis, melanoma,
myeloma, and some cancers. Antibodies against HGFH2 and antisense
sequences may also be used as antagonist for treating the above.
CC Developmental disorders treated by HGFH2 include renal tubular acidosis,
anaemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and
Becker muscular dystrophy, and epilepsy. HGFH2 antibodies and HGFH2
polynucleotides and polypeptides may also be used in various diagnostic
methods.
XX
SQ Sequence 3805 BP; 1255 A; 654 C; 730 G; 1165 T; 1 other;

Query Match 95.7%; Score 980; DB 20; Length 3805;
Best Local Similarity 99.1%; Pred. No. 1.5e-251;
Matches 1017; Conservative 0; Mismatches 5; Indels 4; Gaps 3;

QY 1 ACCGCTCGTCCGCGCGCTTGAGCGCCCGGGAGCGCGCAATTGTCGCGCCCGCGG 60
Db |||||
QY 4 ACCGCTCGTCCG-CGCGCTTGAGCGCCCGGGAG-GCGCGCAATTGTCGCGCCCGCGG 61
Db |||||
QY 61 GGGCGGCGCC--TCCCGGCATCTTCGCGGCGACCAAGGACTACAGGAAGGGAGCGGCTG 118
Db |||||
QY 62 GGGCGGCGCATCCCGCAGTCTTCGCGGCGGACCAAGGACTACAGGAAGGGAGCGGCTG 121
Db |||||


```
XX PA (CORI-) CORIXA CORP.
XX PI Benson DR, Lodes MJ, Mitcham JL, King GE;
XX XX WPI; 2003-147101/14.
XX DR
XX PT Determining presence or absence of cancer in patient by contacting
XX PT patient sample with oligonucleotide that hybridises to polynucleotide
XX PT encoding ovarian carcinoma antigen, and detecting amount of
XX PT hybridisation in sample
XX XX
XX PS Example 1; Fig 16I; 152pp; English.
XX CC
XX CC The present invention relates to compositions and methods for
XX CC determining presence or absence of cancer in patient. The method
XX CC comprising contacting a biological sample with an oligonucleotide
XX CC that hybridises to partial polynucleotide sequence encoding for
XX CC human ovarian carcinoma antigen, or its complement. The method and
XX CC compositions are useful for the therapy and diagnosis of ovarian
XX CC cancer. The method can be used to identify tumour antigens that
XX CC are secreted from ovarian carcinoma and/or other tumours. Effective
XX CC cancer detection is achieved using the method of the invention.
XX CC ABX72720-ABX73050 represent partial cDNA sequences encoding human
XX CC ovarian carcinoma antigens.
XX SQ Sequence 1664 BP; 566 A; 280 C; 319 G; 499 T; 0 other;

Query Match 80.9%; Score 828.4; DB 25; Length 1664;
Best Local Similarity 98.7%; Pred. No. 3.6e-211;
Matches 835; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 179 AGGCTACCCGCACTGCGCGCGATTTGATGAATCCAGAGGGCGCTGTGAAGCCTC 238
DB 17 AGGCGCGCCAGTCGCGGGTACCGCGATTTGATGAATCCAGAGGGCGCTGTGAAGCCTC 76

QY 239 CAGCAAAACAAGATCTCTATCTCTTTTGGCACCCATGAATCTGCAATTTCTAGGTCCCA 298
DB 77 CAGCAAAACAAGATCTCTATCTCTTTTGGCACCCATGAATCTGCAATTTCTAGGTCCCA 136

QY 299 AAGACCTTTTCCATATAAGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAGCGAAAG 358
DB 137 AAGACCTTTTCCATATAAGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAGCGAAAG 196

QY 359 GATTTAACGAAGAGTGTGGGAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTACC 418
DB 197 GATTTAACGAAGAGTGTGGGAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTACC 256

QY 419 AGGCAATTTCAGCAACAGAGCTCTTTCAGAAACTGAGGGAGAGGTGGAAATCTGCAGATG 478
DB 257 AGGCAATTTCAGCAACAGAGCTCTTTCAGAAACTGAGGGAGAGGTGGAAATCTGCAGATG 316

QY 479 CAAGCAGTGAAGAGAGGTGATAGATGATAGAGTGAAGAGTGAAGAGCAAAAGAAAGATG 538
DB 317 CAAGCAGTGAAGAGAGGTGATAGATGATAGAGTGAAGAGTGAAGAGCAAAAGAAAGATG 376

QY 539 AAAAAGCAGGCTCAAAACGGAAGATCATATCTCTCAAGAAATCTCTTAAACAGTCCC 598
DB 377 AAAAAGCAGGCTCAAAACGGAAGATCATATCTCTCAAGAAATCTCTTAAACAGTCCC 436

QY 599 GGAATCTCCAGGAGATGAAGATGACAAAGACTGCAAAAGAGAGGAAACAAAGCAGCT 658
DB 437 GGAATCTCCAGGAGATGAAGATGACAAAGACTGCAAAAGAGAGGAAACAAAGCAGCT 496

QY 659 CTGAGGTGAGATGCGGGCAACGACACAAAGAACACAACTTCAGATTCGAGAAACCA 718
DB 497 CTGAGGTGAGATGCGGGCAACGACACAAAGAACACAACTTCAGATTCGAGAAACCA 556

QY 719 GTGAGGGAGCTTACTACCATATATGATGCTGATATTTAGAGNAACCAAGAGGTTA 778
DB 557 GTGAGGGAGCTTACTACCATATATGATGCTGATATTTAGAGNAACCAAGAGGTTA 616

QY 779 TATGTTTGGTGTCTAATATTTCTTGGATTTGATATGAACCAACACATAGTCTTGTGTC 838
```

```
DB 617 TATGTTTGGTGTCTAATATTTCTTGGATTTGATGAACCAACACATAGTCTTGTGTC 676
QY 839 ATTGACAGAACCCAGTTTGTATGATACATTAATTCATCTCTCTGTTGTTGGGG 898
DB 677 ATTGACAGAACCCAGTTTGTATGATACATTAATTCATCTCTCTGTTGTTGGGG 736
QY 899 GGAAGAGACATTTTAGCCTTTTAAAGTTACTGATTAATTTCAATTTTATTGTTGTC 958
DB 737 GGAAGAGACATTTTAGCCTTTTAAAGTTACTGATTAATTTCAATTTTATTGTTGTC 796
QY 959 ATGAAGTTGCCCTTAACCACTAAGGATTAAGAGTTTTCGCGAGACTTATACATGCT 1018
DB 797 ATGAAGTTGCCCTTAACCACTAAGGATTAAGAGTTTTCGCGAGACTTATACATGCT 856
QY 1019 AGGATC 1024
DB 857 AGGATC 862

RESULT 10
ABX72983
ID ABX72983 standard; cDNA; 1664 BP.
XX AC ABX72983;
XX XX
XX DT 14-MAR-2003 (first entry)
XX DE Human ovarian carcinoma antigen partial cDNA sequence #249.
XX KW Human; cancer detection; ovarian carcinoma antigen; ovarian cancer;
XX KW tumour antigen; tumour; OV2; OV3; OV6; OV9; OV10; OV12; OV14; OV17;
XX KW OV18; OV23; OV24; OV27; OV41; OV54; OV57; gene; ss.
XX OS Homo sapiens.
XX XX
XX PN US6468758-B1.
XX PD 22-OCT-2002.
XX XX
XX PF 16-SEP-1999; 99US-0397787.
XX PR 23-SEP-1998; 98US-0159320.
XX PR 08-FEB-1999; 99US-0246429.
XX XX
XX PA (CORI-) CORIXA CORP.
XX XX
XX PI Benson DR, Lodes MJ, Mitcham JL, King GE;
XX DR WPI; 2003-147101/14.
XX XX
XX PT Determining presence or absence of cancer in patient by contacting
XX PT patient sample with oligonucleotide that hybridises to polynucleotide
XX PT encoding ovarian carcinoma antigen, and detecting amount of
XX PT hybridisation in sample
XX XX
XX PS Example 1; Column 193-196; 152pp; English.
XX CC
XX CC The present invention relates to compositions and methods for
XX CC determining presence or absence of cancer in patient. The method
XX CC comprising contacting a biological sample with an oligonucleotide
XX CC that hybridises to partial polynucleotide sequence encoding for
XX CC human ovarian carcinoma antigen, or its complement. The method and
XX CC compositions are useful for the therapy and diagnosis of ovarian
XX CC cancer. The method can be used to identify tumour antigens that
XX CC are secreted from ovarian carcinoma and/or other tumours. Effective
XX CC cancer detection is achieved using the method of the invention.
XX CC ABX72720-ABX73050 represent partial cDNA sequences encoding human
XX CC ovarian carcinoma antigens.
XX SQ Sequence 1664 BP; 566 A; 280 C; 319 G; 499 T; 0 other;

Query Match 80.9%; Score 828.4; DB 25; Length 1664;
```

Best Local Similarity 98.7%; Pred. No. 3.6e-211;
Matches 835; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 179 AGGGCTACCGCACTCCGCGCGCGCGGATGATGAATCCAGAGGGCGCTGTGAAGCCTC 238
Db 17 AGGGCGCGCGAGTGCAGGCTAGCCGCGATGATGAATCCAGAGGGCGCTGTGAAGCCTC 76

QY 239 CAGCAACAGATATCTATCTCTCTTTTGGACCCATGAACTGATTTCTAGTCCCA 298
Db 77 CAGCAACAGATATCTATCTCTCTTTTGGACCCATGAACTGATTTCTAGTCCCA 136

QY 299 AAGACCTTTTCCATATAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACAAACGGAAAG 358
Db 137 AAGACCTTTTCCATATAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACAAACGGAAAG 196

QY 359 GATTAAACAGAGATTGTGGGAAATAGAAAATAAACCCAGGAGTAAAGTTTACTTGGCTACC 418
Db 197 GATTAAACAGAGATTGTGGGAAATAGAAAATAAACCCAGGAGTAAAGTTTACTTGGCTACC 256

QY 419 AGGCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGAGAGGTGGAATATCTGCAGATG 478
Db 257 AGGCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGAGAGGTGGAATATCTGCAGATG 316

QY 479 CAAGCAGTCAGGAAGAGGTGATAGAGTAGAAGATGAAAGGCAAAAGCAAAAGAAATG 538
Db 317 CAAGCAGTCAGGAAGAGGTGATAGAGTAGAAGATGAAAGGCAAAAGCAAAAGAAATG 376

QY 539 AAAACAGGCTCAAAACGGAAAAAGTCAATATCTTCAAGAAATCCTCTAAACAGTCCC 598
Db 377 AAAACAGGCTCAAAACGGAAAAAGTCAATATCTTCAAGAAATCCTCTAAACAGTCCC 436

QY 599 GGAATCTCCAGAGATGAAGTACAAAGACTGCAAGAGAGGAAAGCAAAACAGAGT 658
Db 437 GGAATCTCCAGAGATGAAGTACAAAGACTGCAAGAGAGGAAAGCAAAACAGAGT 496

QY 659 CTGAGGCTGAGATGCGGCGCAACGACAAAGAAACAACTTCAGACTTGCAGAAACCA 718
Db 497 CTGAGGCTGAGATGCGGCGCAACGACAAAGAAACAACTTCAGACTTGCAGAAACCA 556

QY 719 GTGAAGGACCTTAACCTAACATATGAATGCTGCATATTAAGAGAAACCAAGAGGTTA 778
Db 557 GTGAAGGACCTTAACCTAACATATGAATGCTGCATATTAAGAGAAACCAAGAGGTTA 616

QY 779 TATGTTTGTGTCATATCTTGTGATGATGAACCAACACATAGTCTGTGTC 838
Db 617 TATGTTTGTGTCATATCTTGTGATGATGAACCAACACATAGTCTGTGTC 676

QY 839 ATTGACAGAACCCAGTTGTATGATATTAATTCATATTCCTCTGTTGTGTTTGGGG 898
Db 677 ATTGACAGAACCCAGTTGTATGATATTAATTCATATTCCTCTGTTGTGTTTGGGG 736

QY 899 GGAAGAGACATTTAGCTTTTAAAGTTACTGATTAATTTCAATGTTTGTGTTGTC 958
Db 737 GGAAGAGACATTTAGCTTTTAAAGTTACTGATTAATTTCAATGTTTGTGTTGTC 796

QY 959 ATCAAGTCCCTTACCACTAGGATATCAAGATTTTGGCAGACTTATACATGCT 1018
Db 797 ATCAAGTCCCTTACCACTAGGATATCAAGATTTTGGCAGACTTATACATGCT 856

1019 AGGATC 1024
857 AGGATC 862

RESULT 11
ABL35012
ID ABL35012 standard; cDNA; 2920 BP.

XX ABL35012;

AC ABL35012;

XX 04-APR-2002 (first entry)

DE Rat cDNA isolated from skin cells SEQ ID NO: 541.

XX Human; rat; mouse; skin cell; skin wound; cancer; growth defect;
KW developmental defect; inflammatory disease; dermatological; vulnary;
KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective; gene;
XX ss.
OS Rattus sp.
XX WO200190357-A1.
XX 29-NOV-2001.
XX 24-MAY-2001; 2001WO-NZ00099.
XX 24-MAY-2000; 2000US-206650P.
XX 25-JUL-2000; 2000US-221232P.
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG, Kumble KD;
XX WPI; 2002-122020/16.
XX P-PSDB; ABB72327.
XX New polynucleotides and polypeptides encoded by the polynucleotides
XX isolated from skin cells, useful for treating skin wounds, cancers,
XX growth and developmental defects, inflammatory diseases, or for
XX modulating immune responses -
XX Claim 1; Page 334-335; 466pp; English.
XX The present invention provides the protein and coding sequences of cDNAs
XX isolated from human, murine and rat skin cell libraries. The sequences
XX can be used in the development of therapeutic agents useful in the
XX treatment of skin diseases, including skin wounds, cancer, growth
XX defects, developmental defects and inflammatory diseases. The proteins
XX have important roles in the induction of hair growth, cell proliferation
XX and cell-cell interaction, in maintaining tissue integrity, in wound
XX healing and in modulating immune responses. The present sequence is a
XX cDNA of the invention.
SQ Sequence 2920 BP; 877 A; 600 C; 652 G; 791 T; 0 other;

Query Match 76.0%; Score 778; DB 24; Length 2920;
Best Local Similarity 89.1%; Pred. No. 1.4e-197;
Matches 912; Conservative 0; Mismatches 95; Indels 17; Gaps 6;

QY 2 CCGCTCGTCCGCGCGCGCTTGAGGCGCGCGGAGCGC-GCGCAATTCGTTCGCGCGCGG 60
Db 164 CCGCTCGAGCGCGCGCTCGAGGCGCGCGGAGCGCGCGCGTCCGTTCGTCGCGGA 223

QY 61 GGGCGGCGCTCCCGGATCTTCGCGCGACCAAGACTACCAAGAGGGAGCGGCTGG 120
Db 224 GGGCGGCGCTCCCGGATCTTCGCGCGACCAAGACTACCAAGAGGGAGCGGCTGG 283

QY 121 ATGCGCGCTCCGCGCGCGCGCGAGTACAAAGCGGCGACCTGCTTCGCAAGATGAAG 180
Db 284 ATGCGCGCTCCGCGCGCGCGCGAGTACAAAGCGGCGAGCTGCTTCGCAAGATGAAG 343

QY 181 GGCTACCGCGCACTGGCGCGCGCGGATTCGATGAACCTCCAGAGGGCGCTGTGAAGCTCCA 240
Db 344 GGCTACCGCGCACTGGCGCGCGCGGATTCGATGAACCTCCAGAGGGCGCTGTGAAGCTCCA 403

QY 241 GCAAAACAGTATCTTATCTTTTGGGACCCATGAAGTGAATTTCTAGGTCCTCAA 300
Db 404 GCAAAACAGTATCTTATCTTTTGGGACCCATGAAGTGAATTTCTAGGTCCTCAA 463

QY 301 GACCTTTTCCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAGA 360
Db 464 GACCTTTTCCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAGA 523

QY 361 TTTAACGAGGATTTGTGGGAAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTACCAG 420

Db 524 TTAAATGAAGATTATGGGAAATTGAAATAATCCAGAGTGAAATTTACTGGGTACCAG 583
Qy 421 GCAATTCAGCAACAGAGCTCTTCAGAAATCGAGGAGAGGTGGAAATACTCGACATGCA 480
PI 584 ACAATTCAGCAACAGAGCTCTTCAGAAATCGAGGAGAGGTGGAAATACTCGACATGCA 643
Qy 481 AGCAGTCAGGAAGAGTGTAGAGTAGAGAGATGGAAGGCAAAAGAAAGATGAA 540
Db 644 AGCAGTCAGGAAGAGTGTAGAGTAGAGATGGAAGGCAAAAGAAAGATGAA 700
Qy 541 AAAGCAGGCTCAAAACCGGAAAGTGTATATCTTCAAGAGAAATCTCTAACAGTCCCGG 600
Db 701 AAAGCAGGCTCAAAACCGGAAAGTGTATATCTTCAAGAGAAATCTCTAACAGTCCCGG 760
Qy 601 AAATCTCCAGGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAG 660
Db 761 AAATCTCCAGGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAG 820
Qy 661 GAGGTGGAGATGCGGCAACGACACAAAGAAACAACTTCAGACTTGCAGAAACCAAGT 720
Db 821 GAGGTGGAGATGCGGCAACGACACAAAGAAACAACTTCAGACTTGCAGAAACCAAGT 880
Qy 721 GAAGGACCTAACTACCATATGATGCTGCATATTAAGAGAACCAACCAAGAGTTATA 780
Db 881 GAAGGACCTAACTACCATATGATGCTGCATATTAAGAGAACCAACCAAGAGTTATA 940
Qy 781 TGTTGGTGTCTTAATATCTTGGATTTGATATGAACCAACACATAGTCTTGTGTGCAT 840
Db 941 TGTTGGTGTCTT--GATCTTGGATTTGATATGAACCAAC--AGTCTTGTGTGCAT 994
Qy 841 TGACAGAACCCAGTTTGTATGATATTAATTCATATTCCTCTCTGTGTGTGTGTGTGTGT 900
Db 995 TGACAGAACCCAGTTTGTATGATATTAATTCATATTCCTCTCTGTGTGTGTGTGTGTGT 1050
Qy 901 AAAGACATTTTAGCTTTTAAAGTTACTGATTTAAATTCATGTTATTTGGTGTGCAT 960
Db 1051 AAAGACATTTTAGCTTTTAAAGTTACTGATTTAAATTCATGTTATTTGGTGTGCAT 1110
Qy 961 GAAGTTGCTCCCTTAACCACTAAGGATTATCAAGATTTTGGCGACATTAACATGCTAG 1020
Db 1111 GAAGTTGCTCCCTTAACCACTAAGGATTATCAAGATTTTGGCGACATTTTGGCGACAT 1167
Qy 1021 GATC 1024
Db 1168 GCTC 1171

RESULT 12
AAH05630
ID AAH05630 standard; cDNA; 737 BP.
XX
AC AAH05630;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA clone (5'-primer) SEQ ID NO:2465.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
OS
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX

PA (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -

XX Claim 1; SEQ ID 2465; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH98893 represent human amino acid sequences; and AAH13629 to
CC AAH13632 represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 737 BP; 227 A; 180 C; 204 G; 117 T; 9 other;

Query Match 60.9%; Score 623.6; DB 22; Length 737;
Best Local Similarity 96.2%; Pred. No. 1.5e-156;
Matches 676; Conservative 0; Mismatches 23; Indels 4; Gaps 4;

Qy 2 CGCTCGTCGCGCCGCGCTTTGAGCGCCGCGGGAGCGC-GCGCAATTCGTGCGGCCCGCGG 60

Db 36 CGCTCGTCGCGCCGCGCTTTGAGCGCCGCGGGAGCGC-GCGCAATTCGTGCGGCCCGCGG 95

Qy 61 GGGCGCGCTCCCGGCATCTTCGCGCGCAACCAAGGACTACCAAGGAGGGGAGCGGTGGG 120

Db 96 GGGCGCGCTCCCGGCATCTTCGCGCGCAACCAAGGACTACCAAGGAGGGGAGCGGTGGG 155

Qy 121 ATGGCGCGTCGCGGCCCGCGAGTACAAAGCGGGGAGCTGGTCTTCGCAAGATGAAG 180

Db 156 ATGGCGCGTCGCGGCCCGCGAGTACAAAGCGGGGAGCTGGTCTTCGCAAGATGAAG 215

Qy 181 GGTACCCGCTGCGCGGCCCGCGAGTACAAAGCGGGGAGCTGGTCTTCGCAAGATGAAG 240

Db 216 GGTACCCGCTGCGCGGCCCGCGAGTACAAAGCGGGGAGCTGGTCTTCGCAAGATGAAG 275

Qy 241 GCAAAACAAGTATCTATCTTTTGGCACCCCATGAAGTGCATTTCTAGTCCCAA 300

Db 276 GCAAAACAAGTATCTATCTTTTGGCACCCCATGAAGTGCATTTCTAGTCCCAA 335

Qy 301 GACCTTTTCCATATAAGAGTACAAAGCAAGTTTGGAAAGTCAAAACAACGGAAGGA 360

Db 336 GACCTTTTCCATATAAGAGTACAAAGCAAGTTTGGAAAGTCAAAACAACGGAAGGA 395

Qy 361 TTTAACGAAGGATTTGGGAAATAGAAATAACCCAGGAGTAAAGTTTACTGGCTACCAG 420

Db 396 TTTAACGAAGGATTTGGGAAATAGAAATAACCCAGGAGTAAAGTTTACTGGCTACCAG 455

Qy 421 GCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAAAGTGGAAATATCTGCAGATGCA 480

Db 456 GCAATTTCAGCAACAGAGCTCTTTCAGAACTGAGGAGAGAAAGTGGAAATATCTGCAGATGCA 515
Qy 481 AGCAGTGTAGGAAGAGGTGATAGACTAGAGAGAGATGGAAGAGGCAAAAGAAAGATGAA 540
Db 516 AGCAGTGTAGGAAGAGGTGATAGACTAGAGAGAGATGGAAGAGGCAAAAGAAAGATGAA 575
Qy 541 AAAGCAGGCTCAAAACGGAAGAGTATATATCTTCAAAAGAAATCTCTAAACAGTCCCGG 600
Db 576 AAAGCAGGCTCAAAACGGAAGAA-TCATATATCTTCAAGAAATCTCTAAACAGTCCCGG 634
Qy 601 AAATCTCCAGG-AGATGAGATGACAAAGACTGCAAGAGAGAGGAAACAAAGAGAGTCTC 659
Db 635 AAATCTCCAGGAGATGAGATGACAAAGACTGCCNAGAAAGAGAAACNAAAGCAGTCTC 694
Qy 660 TGAGGCTGAGATG-CGGGCAACGACACAGAAACACACTTC 701
Db 695 TGAAGGTGGANATGCCGGCAACGACACNAGAAACACACTTC 737

RESULT 13
AA65004
ID AA65004 standard; cDNA; 609 BP.
XX AC AA65004;
XX DT 13-NOV-2000 (first entry)
XX DE Human hepatoma-derived growth factor homologous polypeptide cDNA.
XX KW Human; hepatoma-derived growth factor homologous polypeptide;
KW huHDGFh; cell proliferation; haematopoiesis; angiogenesis;
KW lymphocyte proliferation; infection; autoimmune disease;
KW vascular disease; cancer; ss.
XX OS Homo sapiens.

XX FH Key Location/Qualifiers
XX CDS 1..609
FT /*tag= a
FT /product= "huHDGFh"
XX PN WO200037492-A2.
XX PD 29-JUN-2000.
XX PF 22-DEC-1999; 99WO-US30932.
XX PR 22-DEC-1998; 98US-0113344.
XX PA (ELIL) LILLY & CO ELI.
XX PI Na S;
XX DR WPI; 2000-442638/38.
XX P-PSDB; AAB13521.
XX PT Novel human homologues of hepatoma-derived growth factor homologous
PT nucleic acids, polypeptides useful as probes or amplification primers
PT in the detection, quantitation or isolation of gene sequences or
PT transcripts
XX PS Claim 3; Page 73-74; 75pp; English.
XX CC The present sequence is the human homologue of the hepatoma-derived
CC growth factor coding sequence (huHDGFh). The protein produced from this
CC gene is involved in tumour formation in some cells, and it is thought
CC that the gene and protein will be useful in the diagnosis and treatment
CC of infections, autoimmune disorders, vascular diseases and cancers. In
CC addition, the gene can be used to produce transgenic animals which may be
CC used as animal models for these diseases, and the protein can be used to
CC induce cell proliferation, haematopoiesis, lymphocyte proliferation and
CC angiogenesis.
XX

SQ Sequence 609 BP; 223 A; 125 C; 159 G; 102 T; 0 other;
Query Match 59.5%; Score 609; DB 21; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.1e-152;
Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 121 ATGGCGGTCCGCGGCCCGCGAGTACAAAGCGGGCGACCTTGCTTTCGCAAGATGAAG 180
Db 1 ATGGCGGTCCGCGGCCCGCGAGTACAAAGCGGGCGACCTTGCTTTCGCAAGATGAAG 60
Qy 181 GGTACCCGCACTGGCGGCCCGCGAGTATGATGAATCTCCAGAGGGCGCTGTGAAGCTTCA 240
Db 61 GGTACCCGCACTGGCGGCCCGCGAGTATGATGAATCTCCAGAGGGCGCTGTGAAGCTTCA 120
Qy 241 GCAACAGATATCTATCTCTCTTTTGGGCAACCATGAACTGCAATTTCTAGTCTCCAAA 300
Db 121 GCAACAGATATCTATCTCTCTTTTGGGCAACCATGAACTGCAATTTCTAGTCTCCAAA 180
Qy 301 GACCTTTTTCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACCGGAAAGGA 360
Db 181 GACCTTTTTCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACCGGAAAGGA 240
Qy 361 TTTAACGAGATTTGTGGGAAATAGAAATTAACCCAGGAGTAAAGTTTACTTGGCTACCAG 420
Db 241 TTTAACGAGATTTGTGGGAAATAGAAATTAACCCAGGAGTAAAGTTTACTTGGCTACCAG 300
Qy 421 GCAATTCAGCAACAGAGCTCTTCAGAACTCAGGAGAGAGTGGAAATCTGCAGATGCA 480
Db 301 GCAATTCAGCAACAGAGCTCTTCAGAACTCAGGAGAGAGTGGAAATCTGCAGATGCA 360
Qy 481 AGCAGTGTAGGAAGAGGTGATAGAGTAGAAGAGATGGAAGAGGCAAAAGAAAGATGAA 540
Db 361 AGCAGTGTAGGAAGAGGTGATAGAGTAGAAGAGATGGAAGAGGCAAAAGAAAGATGAA 420
Qy 541 AAAGCAGGCTCAAAACGGAAGAGTCTATATATCTTCAAGAAATCTCTTAAACAGTCCCGG 600
Db 421 AAAGCAGGCTCAAAACGGAAGAGTCTATATATCTTCAAGAAATCTCTTAAACAGTCCCGG 480
Qy 601 AAATCTCCAGGAGATGAGATGACAAAGACTGCAAGAGAGAGGAAACAAAGAGAGTCT 660
Db 481 AAATCTCCAGGAGATGAGATGACAAAGACTGCAAGAGAGAGGAAACAAAGAGAGTCT 540
Qy 661 GAGGTGTGAGATGGCGGCAACGACACACAGACACACACTTTCAGACTTGCAGAAACCCAGT 720
Db 541 GAGGTGTGAGATGGCGGCAACGACACACAGACACACACTTTCAGACTTGCAGAAACCCAGT 600
Qy 721 GAAGGGACC 729
Db 601 GAAGGGACC 609
RESULT 14
AA679520
ID AA679520 standard; cDNA; 870 BP.
XX AC AA679520;
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #15324.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.

```

PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
PA
XX Dmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG15333.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 1; SEQ ID No 15324; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAG64197-AAG94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 870 BP; 277 A; 187 C; 256 G; 150 T; 0 other;
XX
XX Query Match 54.2%; Score 555.4; DB 23; Length 870;
XX Best Local Similarity 91.1%; Pred. No. 2.8e-138;
XX Matches 746; Conservative 0; Mismatches 56; Indels 17; Gaps 14;
XX
XX 16 GCGTTCGAGCGCGCGGAGCGCGCAATTCTCGCGCGCGCGGGGGCGGCTCCCGG 75
XX |||||
XX 52 GCGCGGGGGCGCATAGCCCGGCAATACGCGCCCGGCGGCGGGCGGCTCCCGG 111
XX |||||
XX 76 CAT-CTTCGCGCGCAAGGACTACAGGAAGGGGAGCGGCTGGGATGCGCGTCCCGG 134
XX |||||
XX 112 CATGCTTCGCGCGCAAGGACTACAGGAAGGGGAGCGGCTGGGATGCGCGTCCCGG 171
XX |||||
XX 135 GCGCGCGGAG-TACAAGCGGGGACCTGTCTTC-GCCAGATGAAGGGCTACCGGCAC 192
XX |||||
XX 172 GTCCCGCGAGATCAAAAGCGGGGACCTGTGTTCTCGGCCAAGATGAAGGGCTACCGGCAC 231
XX |||||
XX 193 TGCGCGCGCGGATTCATGAACCTCCAGAGCGGCTGTGAAGCCTCCAGCAAC-AAGTA 251
XX |||||
XX 232 TGCGCGCGCGGATTCATGAACCTCCAGAGCGGCTGTGTAAGCCTCAACAAAGTA 291
XX |||||
XX 252 TCCTATCTCTTTTGG---CACCCATGAACCTGCAATTTCTAGGT-CCCAAGACCTTT 307
XX |||||
XX 292 TCCTATCTCTTTTGGCACCCCATGAACCGGCATTTCTAGGTCCCAATGACCTTT 351
XX |||||
XX 308 TTCATATAGGAGTACAAGACAAGTTT-GGAAAGTCAACAAACGGAAGGATTTAAC 366
XX |||||
XX 352 TTCATATAGGAGTACAAGACAAGTTTGGGAAAGTCAACAAACGGAAGGATTTAAC 411
XX |||||
XX 367 GAAGGATTTG-TGGAAATAGAAATAACCCAGG--AGTAAAGTTTACTGGCTACCAAGGC- 422
XX |||||
XX 412 GAAGGATTTGGGGAATAGAAATAACCCAGGAGTAAAGGTTTACTGGCTACCAAGGC 471
XX |||||
XX 423 AATTCACCAACAGAGCTCTTCAGAAAACCTGAGGGAGAGGTGGAAATACTGCAGATGCAAG 482
XX |||||

Db 472 AATTGAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTGGAAATACTGCAGATGCAAG 531
Qy 483 CAGTGAAGGAAGAGGTGATAGAGTGAAGAAGAT-CGAAAGGCAAAAGCAAGAAATGAAA 541
Db 532 CAGTGAAGGAAGAGGTGATAGAGTGAAGAAGATGGGAAAGGCAAAAGCAAGAAATGAAA 591
Qy 542 AAGCAGGCTCAAAAACGGAAGAAAGTTCATATATCTTTCAAGAAATCTCTTAAACAGT-CCCCG 600
Db 592 AAGCAGGCTCAAAAACGGAAGAAAGTTCATATATCTTTCAAGAAATCTCTTAAACAGTCCCCG 651
Qy 601 AATCTCCAGGAGATGAAGA-TGACAAAGACTGCAAAAGAGGAGGAAACAAAAGCAGCTC 659
Db 652 AATCTCCAGGAGATGAAGATTGACAAAGACTGCAAAAGAGGAGGAAACAAAAGCAGCTC 711
Qy 660 TGAGGGTGGAGATG-CGGGCAACGACACAAAGAAACACAACTTCAGACTTCGAGAAACCA 718
Db 712 TGAGGGTGGAGATGCGCGGCAACGACACAAAGAAACACAACTTCAGACTTCGAGAAACCA 771
Qy 719 GTGAAGGGACCTAACTTACCATAATGAATGCTGCATATTAAAGAGAAACCAAGAGGTTA 778
Db 772 GTGAAGGGACCTAACTTACCATAATGAATGCTGCATATTAAAGAGAAACCAAGAGGTTA 831
Qy 779 TATGTTTGGTTGCTTAATATTTCTTGGATTGATATGAAC 817
Db 832 TATGTTTGGTTGCTTAATATTTCTTGGATTGATATGAAC 870

RESULT 15
AAZ80234/c
ID AAZ80234 standard; cDNA; 645 BP.
XX
XX AAZ80234;
XX
XX 07-APR-2000 (first entry)
XX
XX Human colon cancer cell line SW480 cDNA clone SEQ ID NO:318.
XX
XX Human; gene expression product; diagnosis; tumour; colon cancer;
XX colorectal adenocarcinoma; cell line SW480; cell proliferation;
XX cytotatic; sarcoma; breast cancer; neoplasia; dysplasia;
XX
XX Homo sapiens.
XX
XX W09964576-A2.
XX
XX 16-DEC-1999.
XX
XX 09-JUN-1999; 99WO-IB01062.
XX
XX 10-JUN-1998; 98US-0088801.
XX
XX (FARB ) BAYER CORP.
XX
XX Endege WO, Steinmann KE, Astle JH, Burgess CC, Bushnell SE;
XX Carroll E, Catino TJ, Derti A, Ford DM, Lewis ME, Monahan JE;
XX Schlegel R;
XX
XX WPI; 2000-087220/07.
XX
XX Novel nucleic acids, used to develop products for the diagnosis and
XX treatment of disorders involving unwanted cell proliferation,
XX particularly cancers, especially colon cancer -
XX
XX Claim 15; Page 260; 469pp; English.
XX
XX AAZ79917 to AAZ80766 represent double stranded cDNA clones isolated from
XX the human colorectal adenocarcinoma (colon cancer) cell line SW480. The
XX cDNA clones can be used to generate antisense oligonucleotides which
XX can be used for antisense therapy. Methods and products from the present
XX invention can be used for identifying and/or classifying cancerous cells
XX present in a human tumour, particularly in solid tumours, e.g.
```


CC carcinomas and sarcomas, e.g. breast or colon cancers. The cDNA clones
CC can be used for developing agents for the diagnosis and treatment of
CC disorders involving unwanted cell proliferation, such as neoplasia,
CC dysplasia or hyperplasia.

```
XX
SQ Sequence 645 BP; 123 A; 160 C; 123 G; 233 T; 6 other;
Query Match          46.0%; Score 471; DB 21; Length 645;
Best Local Similarity 95.8%; Pred. No. 9.2e-116;
Matches 525; Conservative 0; Mismatches 18; Indels 5; Gaps 4;
QY 320 AGTACAAAGACAAAGTTTGAAGTCAAAACAAACGGAAGGATTTAAACGAAGGATTGTGGG 379
Db 557 AGGTCACANGCCAGTTTGAAGTC-AACCAACGGAAGGATTTAAACGAAGGATTGTGGA 499
QY 380 AAATAGAAATTAACCCAGAGATAAGTTTACTGGCTACCAAGCAATTCAGCAACAGAGCT 439
Db 498 AATAGAAAAT--ACCCAGNAGTAAAGTTTACTGG-TACCAGGCAATTCAGCACAGAGNT 442
QY 440 CTTCAAGAACTGAGGAGAGGAGTGAATATCTGCAGATGCAAGCAGTGTAGGAGAGGTTG 499
Db 441 CTTCAAGAACTGA-GGAGAGGTTGGAATATCTGCAGATGCAAGCAGTGTAGGAGAGGTTG 383
QY 500 ATAGAGTAGAAGAGATGGAAGAGCAAAAGGCAAAAGAAAGAAATCAAAAACGAGCTCAAAAACGGA 559
Db 382 ATAGAGTAGAAGAGATGGAAGAGCAAAAGGCAAAAGAAAGAAATCAAAAACGAGCTCAAAAACGGA 323
QY 560 AAAAGTCATATCTTCAAGAAATCTCTTAAACAGTCCCGGAAATCTCCAGGAGATGAAG 619
Db 322 AAAAGTCATATCTTCAAGAAATCTCTTAAACAGTCCCGGAAATCTCCAGGAGATGAAG 263
QY 620 ATGACAAAGACTGCAAGAGAGGAAACAAAGCAGCTCTGAGGGTGGAGATCGGGCA 679
Db 262 ATGACAAAGACTGCAAGAGAGGAAACAAAGCAGCTCTGAGGGTGGAGATCGGGCA 203
QY 680 ACGCACAAAGAACACAACTTCAGACTTGCAGAAACCAAGTGAAGGGACCTAACTACCAT 739
Db 202 ACGCACAAAGAACACAACTTCAGACTTGCAGAAACCAAGTGAAGGGACCTAACTACCAT 143
QY 740 AATGAATGCTGCATATTAAGAAACCAAGAGTTATATGTTGGTTGTCTAATATT 799
Db 142 AATGAATGCTGCATATTAAGAAACCAAGAGTTATATGTTGGTTGTCTAATATT 83
QY 800 CTTGGATTGTGATATGAACCAACATAGTCTTGTTCATTGACAGAACCCAGTTTGT 859
Db 82 CTTGGATTGTGATATGAACCAACATAGTCTTGTTCATTGACAGAACCCAGTTTGT 23
QY 860 ATGTACAT 867
Db 22 ATGTACCT 15
```

Search completed: January 18, 2004, 09:11:01
Job time : 332 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2004, 08:55:40 ; Search time 78 Seconds
(without alignments)
5794.568 Million cell updates/sec

Title: US-09-787-328B-3
Perfect score: 1024
Sequence: 1 accgctcgccgccggctt.....actatacatgtctaggatc 1024

Scoring table: IDENTITY NUC
Gapex 10.0, Gapex 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	969	94.6	4487	US-09-484-970B-96	Sequence 96, Appl
2	828.4	80.9	1664	US-09-397-787-68	Sequence 68, Appl
3	828.4	80.9	1664	US-09-397-787-267	Sequence 267, Appl
C 4	471	46.0	645	US-09-328-111-318	Sequence 318, App
5	193.4	18.9	2376	US-08-760-745-4	Sequence 4, Appli
6	188.6	18.4	869	US-08-760-745-2	Sequence 2, Appli
C 7	98.2	9.6	7218	US-08-232-463-14	Sequence 14, Appl
8	50	4.9	696	US-09-461-697-193	Sequence 193, App
9	50	4.9	699	US-09-461-697-191	Sequence 191, App
10	50	4.9	717	US-09-461-697-189	Sequence 189, App
11	50	4.9	774	US-09-461-697-187	Sequence 187, App
12	50	4.9	819	US-09-461-697-185	Sequence 185, App
13	50	4.9	1280	US-09-220-132-155	Sequence 155, App
14	50	4.9	1669	US-09-461-697-184	Sequence 184, App
15	47.6	4.6	2223	US-08-257-073-4	Sequence 4, Appli
C 16	47.2	4.6	2791	US-09-570-367C-1	Sequence 1, Appli
17	46.8	4.6	1298	US-08-971-089-1	Sequence 1, Appli
18	45.4	4.4	1194	US-08-929-329-3	Sequence 3, Appli
19	45.2	4.4	340	US-08-182-175A-104	Sequence 104, App
20	45.2	4.4	340	PCT-US92-06412-104	Sequence 104, App
C 21	45.2	4.4	929	US-09-671-317-14	Sequence 14, Appl
C 22	45.2	4.4	1001	US-09-671-317-439	Sequence 439, App
C 23	44	4.3	4453	US-09-146-053-5	Sequence 5, Appli
24	43.8	4.3	289	US-09-007-005-17	Sequence 17, Appl
25	43.8	4.3	289	US-09-244-796-17	Sequence 17, Appl
26	42.6	4.2	1298	US-08-948-705-3	Sequence 3, Appli
27	42.6	4.2	1298	US-09-510-543-3	Sequence 3, Appli

28	42.6	4.2	6755	3	US-08-931-999-4	Sequence 4, Appli
29	42.4	4.1	282	3	US-09-461-697-205	Sequence 205, App
30	42.4	4.1	306	3	US-09-461-697-203	Sequence 203, App
31	42	4.1	43795	3	US-08-742-185-101	Sequence 101, App
32	41.8	4.1	2447	2	US-09-014-969-14	Sequence 14, Appl
33	41.6	4.1	376	2	US-08-623-906A-18	Sequence 18, Appl
34	41.6	4.1	4226	4	US-09-620-312D-480	Sequence 480, App
35	41.6	4.1	4766	5	PCT-US93-07261-10	Sequence 10, Appl
36	41.6	4.1	6002	4	US-09-345-882-4	Sequence 4, Appli
37	41.4	4.0	189	3	US-09-461-697-213	Sequence 213, App
38	41.4	4.0	195	3	US-09-461-697-211	Sequence 211, App
39	41.4	4.0	213	3	US-09-461-697-209	Sequence 209, App
40	41.4	4.0	231	3	US-09-461-697-207	Sequence 207, App
C 41	41.4	4.0	152331	3	US-09-128-155-16	Sequence 16, Appl
42	41.4	4.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
43	40.6	4.0	980	4	US-09-171-209-8	Sequence 8, Appli
44	40.6	4.0	5394	3	US-08-688-376-1	Sequence 1, Appli
45	40.4	3.9	1447	4	US-09-443-041A-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-09-484-970B-96
; Sequence 96, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 96
; LENGTH: 4487
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 330852.1CB1
; NAME/KEY: unsure
; LOCATION: 1985, 1992, 1997
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-96

Query Match 94.6%; Score 969; DB 4; Length 4487;
Best Local Similarity 99.0%; Pred. No. 1.7e-270;
Matches 1017; Conservative 0; Mismatches 5; Indels 5; Gaps 4;
Qy 1 ACCGCTCGTCCGCCCGCTTTAGCGCCCGGGGAGCGCGCAATTGTCGCGCCCGCGG 60
Db 4 ACCGCTCGTCCG-CGCGCTTTCGCGCGCGGAG-GCGCGCAATTGTCGCGCCCGCGG 61
Qy 61 GGGGCGGCC--TCCCGGCATCTTCGCGCGGACCAAGACTACCAAGAGGGGAGCGGCTG 118
Db 62 GGGGCGGCCATCCCGCGAGTCTTCGCGCGGACCAAGACTACCAAGAGGGGAGCGGCTG 121
Qy 119 GGATGGCGCTCGCGCGCCCGGAGTACAAAGCGGCGACCTGCTTCGCGCAAGATGA 178
Db 122 GGATGGCGCTCGCGCGCCCGCGAGTACAAAGCGGCGACCTGCTTCGCGCAAGATGA 181
Qy 179 AGGGCTACCGCGACTCGCGCGCCCGGATGATGAATCCAGAGGGCGCTGTGAAGCCTC 238
Db 182 AGGGCTACCGCGACTCGCGCGCCCGGATGATGAATCCAGAGGGCGCTGTGAAGCCTC 241
Qy 239 CAGCAACAAGTATCTCTTTTGGGACCCATGAATGATGATTTCTAGTGGTCCCA 298
Db 242 CAGCAACAAGTATCTCTTTTGGGACCCATGAATGATGATTTCTAGTGGTCCCA 301

QY 299 AAGACCTTTTTCATATAAGAGTACAAAGACAAAGTTTGGAAAGTCAAAACAAACGGAAG 358
Db 302 AAGACCTTTTTCATATAAGAGTACAAAGACAAAGTTTGGAAAGTCAAAACAAACGGAAG 361
QY 359 GATTTAACGAGGATTTGGGAAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTACC 418
Db 362 GATTTAACGAGGATTTGGGAAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTACC 421
QY 419 AGCAATTCAGCAACAGAGCTCTTCAGAACTCAGGAGAGAGTGAATATCTGCGAGATG 478
Db 422 AGCAATTCAGCAACAGAGCTCTTCAGAACTCAGGAGAGAGTGAATATCTGCGAGATG 481
QY 479 CRAAGAGTGGAGAGAGTGTATAGATGAGAGAGTGAAGAGTGAAGAGGCAAAAGAAAGT 538
Db 482 CRAAGAGTGGAGAGAGTGTATAGATGAGAGAGTGAAGAGTGAAGAGGCAAAAGAAAGT 541
QY 539 AAAAGCAGGCTCAAAACGGAAGAGTGTATATCTTCAAGAGAGTCTTAAACAGTCCC 598
Db 542 AAAAGCAGGCTCAAAACGGAAGAGTGTATATCTTCAAGAGAGTCTTAAACAGTCCC 601
QY 599 GGAATCTCAGGAGATGAAGTACAAAGAGTGCAGAGAGAGGAGGAGGAGGAGGAGT 658
Db 602 GGAATCTCAGGAGATGAAGTACAAAGAGTGCAGAGAGGAGGAGGAGGAGGAGT 661
QY 659 CTGAGGAGTGGAGATGGGCAACGACACAGAGAGTGCAGAGTGCAGAGAGGAGGAGT 718
Db 662 CTGAGGAGTGGAGATGGGCAACGACACAGAGAGTGCAGAGTGCAGAGAGGAGGAGT 721
QY 719 GTGAGGAGCTTAACATATATGATGATGATGATGATGATGATGATGATGATGATGATG 777
Db 722 GTGAGGAGCTTAACATATATGATGATGATGATGATGATGATGATGATGATGATGATG 781
QY 778 ATATGTTTGGTGTCTTAATATCTTGGATTTGATGATGATGATGATGATGATGATGATG 837
Db 782 ATATGTTTGGTGTCTTAATATCTTGGATTTGATGATGATGATGATGATGATGATGATG 841
QY 838 CATGACAGAACCCAGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 897
Db 842 CATGACAGAACCCAGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 901
QY 898 GGGAAAGACATTTTACGCTTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAA 957
Db 902 GGGAAAGACATTTTACGCTTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAA 961
QY 958 CATGAAGTTGCCCTTAACACATGATGATGATGATGATGATGATGATGATGATGATGATG 1017
Db 962 CATGAAGTTGCCCTTAACACATGATGATGATGATGATGATGATGATGATGATGATGATG 1021
QY 1018 TAGGATC 1024
Db 1022 TAGGATC 1028

RESULT 2

US-09-397-787-68
; Sequence 68, Application US/09397787
; Patent No. 6468758

GENERAL INFORMATION:

; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; FILE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 68
; LENGTH: 1664
; TYPE: DNA
; ORGANISM: Homo sapien

RESULT 3

US-09-397-787-267
; Sequence 267, Application US/09397787
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.

US-09-397-787-68
Query Match 80.9%; Score 828.4; DB 4; Length 1664;
Best Local Similarity 98.7%; Pred. No. 5.1e-230;
Matches 835; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 179 AGGGCTACCCCGCAGTGGCGGCGCCGGATTTGATGAACTCCAGAGGCGCTGTGAAGCCTC 238
Db 17 AGGGCGGCGCAGTGGCGGCGCCGGATTTGATGAACTCCAGAGGCGCTGTGAAGCCTC 76
QY 239 CAGCAAAACAGTATCTATCTTTTGGCACCAGTGAATCTGCATTTCTAGGTCCCA 298
Db 77 CAGCAAAACAGTATCTATCTTTTGGCACCAGTGAATCTGCATTTCTAGGTCCCA 136
QY 299 AAGACCTTTTTCATATAAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACAAACGGAAG 358
Db 137 AAGACCTTTTTCATATAAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACAAACGGAAG 196
QY 359 GATTTAACGAGGATTTGGGAAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTACC 418
Db 197 GATTTAACGAGGATTTGGGAAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTACC 256
QY 419 AGGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGAGTGAATATCTGCGAGATG 478
Db 257 AGGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGAGTGAATATCTGCGAGATG 316
QY 479 CAAGCAGTGGAGAGAGTGTATAGATGAGAGAGTGAAGAGTGAAGAGGCAAAAGAAAGT 538
Db 317 CAAGCAGTGGAGAGAGTGTATAGATGAGAGAGTGAAGAGTGAAGAGGCAAAAGAAAGT 376
QY 539 AAAAGCAGGCTCAAAACGGAAGAGTGTATATCTTCAAGAGAGTCTTAAACAGTCCC 598
Db 377 AAAAGCAGGCTCAAAACGGAAGAGTGTATATCTTCAAGAGAGTCTTAAACAGTCCC 436
QY 599 GGAATCTCAGGAGATGAAGTACAAAGAGTGCAGAGTGAAGAGAGGAGGAGGAGT 658
Db 437 GGAATCTCAGGAGATGAAGTACAAAGAGTGCAGAGTGAAGAGGAGGAGGAGT 496
QY 659 CTGAGGAGTGGAGATGGGCAACGACACAGAGAGTGAAGAGTGAAGAGGAGGAGTGA 718
Db 497 CTGAGGAGTGGAGATGGGCAACGACACAGAGAGTGAAGAGTGAAGAGGAGGAGTGA 556
QY 719 GTGAGGAGCCTTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 778
Db 557 GTGAGGAGCCTTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 616
QY 779 TATGTTTGGTGTCTTAATATCTTGGATTTGATGATGATGATGATGATGATGATGATG 838
Db 617 TATGTTTGGTGTCTTAATATCTTGGATTTGATGATGATGATGATGATGATGATGATG 676
QY 839 ATTGACAGAACCCAGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 898
Db 677 ATTGACAGAACCCAGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 736
QY 899 GGAAGAGACATTTTAGCCTTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAA 958
Db 737 GGAAGAGACATTTTAGCCTTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAA 796
QY 959 ATGAAGTTGCCCTTAACACATGATGATGATGATGATGATGATGATGATGATGATGATG 1018
Db 797 ATGAAGTTGCCCTTAACACATGATGATGATGATGATGATGATGATGATGATGATGATG 856
QY 1019 AGGATC 1024
Db 857 AGGATC 862

APPLICANT: Lodes, Michael J.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: King, Gordon E.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
FILE REFERENCE: 210121.466C2
CURRENT APPLICATION NUMBER: US/09/397,787
CURRENT FILING DATE: 1999-09-16
NUMBER OF SEQ ID NOS: 334
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 267
LENGTH: 1664
TYPE: DNA
ORGANISM: Homo sapien
US-09-397-787-267

Query Match 80.9%; Score 828.4; DB 4; Length 1664;
Best Local Similarity 98.7%; Pred. No. 5.1e-230;
Matches 835; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 179 AGGGCTACCGCAGCTGCGCGCGCGGATGATGAATCCCGAGAGGGCGCTGTGAAGCCTC 238
DB 17 AGGGCGCGCCAGTCGCGGTAGCCGGATTGATGNACTCCCGAGAGGGCGCTGTGAAGCCTC 76
QY 239 CAGCAAAACAAGTATCTTCTCTTTTGGCACCCCATGAATGCAATTTCTAGGTCCCA 298
DB 77 CAGCAAAACAAGTATCTTCTCTTTTGGCACCCCATGAATGCAATTTCTAGGTCCCA 136
QY 299 AAGACCTTTTCCATATAAGAGGTACAAAGACAAGTTTGGAAAGTCAAAACGGAAG 358
DB 137 AAGACCTTTTCCATATAAGAGGTACAAAGACAAGTTTGGAAAGTCAAAACGGAAG 196
QY 359 GATTTAAGCAAGAGTGTGGGAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTACC 418
DB 197 GATTTAAGCAAGAGTGTGGGAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTACC 256
QY 419 AGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGGAGAGGTGGAAATCTGCAGATG 478
DB 257 AGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGGAGAGGTGGAAATCTGCAGATG 316
QY 479 CAAGCAGTGAAGAGAGGTGATAGTAGTAAGAGAGTGAAGAGGCAAAAGAAAGATG 538
DB 317 CAAGCAGTGAAGAGAGGTGATAGTAGTAAGAGAGTGAAGAGGCAAAAGAAAGATG 376
QY 539 AAAAAGCAGCTCAAAACGGAAGATCATATCTTCAAGAGAAATCTCTAAACAGTCCC 598
DB 377 AAAAAGCAGCTCAAAACGGAAGATCATATCTTCAAGAGAAATCTCTAAACAGTCCC 436
QY 599 GGAATCTCCAGGAGTGAAGATGACAAAGATGCAAAAGAGGAGGAAACAAAGAGCAGCT 658
DB 437 GGAATCTCCAGGAGTGAAGATGACAAAGATGCAAAAGAGGAGGAAACAAAGAGCAGCT 496
QY 659 CTGAGGTGAGATGCGGGCAACGACACAGAAACACAACTTCAGACTTCGAAATCCGAAACCA 718
DB 497 CTGAGGTGAGATGCGGGCAACGACACAGAAACACAACTTCAGACTTCGAAATCCGAAACCA 556
QY 719 GTGAAGGAGCTTAATCACTAATATGATGCTGATATTAAGAGAAACCAAGAGGTTA 778
DB 557 GTGAAGGAGCTTAATCACTAATATGATGCTGATATTAAGAGAAACCAAGAGGTTA 616
QY 779 TATGTTTGGTGTCTAATATCTTGGATTGATATGAACCAACACAPAGTCTTTGTTGTC 838
DB 617 TATGTTTGGTGTCTAATATCTTGGATTGATATGAACCAACACAPAGTCTTTGTTGTC 676
QY 839 ATTGACAGAACCCAGTTGTATGATCAATTAATCAATTCCTCTCTGTTGTGTTGCGGG 898
DB 677 ATTGACAGAACCCAGTTGTATGATCAATTAATCAATTCCTCTCTGTTGTGTTGCGGG 736
QY 899 GGAAGAGACATTTAGCCTTTTAAAGTTACTGATTAATTTCAATGTTATTTGTTGTC 958
DB 737 GGAAGAGACATTTAGCCTTTTAAAGTTACTGATTAATTTCAATGTTATTTGTTGTC 796
QY 959 ATGAAGTTGCGCTTAACCACTAAGGATTATCAAGATTTTTCGCGCAGACTTATACATGCT 1018

DB 797 ATGAAGTTGCGCTTAACCACTAAGGATTATCAAGATTTTTCGCGCAGACTTATACATGCT 856
QY 1019 AGGATC 1024
DB 857 AGGATC 862.
RESULT 4
US-09-328-111-318/c
Sequence 318, Application US/09328111
Patent No. 6262333
GENERAL INFORMATION:
APPLICANT: Endege, Wilson O.
APPLICANT: Steinmann, Kathleen E.
APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Catino, Theodore J.
APPLICANT: Derti, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TITLE OF INVENTION: PRODUCTS
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/328,111
CURRENT FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: US 60/088,801
EARLIER FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 318
LENGTH: 645
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(645)
OTHER INFORMATION: n = A,T,C or G
US-09-328-111-318

Query Match 46.0%; Score 471; DB 3; Length 645;
Best Local Similarity 95.8%; Pred. No. 1.1e-126;
Matches 525; Conservative 0; Mismatches 18; Indels 5; Gaps 4;
QY 320 AGTACAAAGACAGTTTGGAAAGTCAAAACGGAAGGATTTAAACGAGGATTTGTGG 379
DB 557 AGGTCCAAGCCAGTTTGAAGAGTC-AACCAACGGAAGGATTTAAGGAGGATTTGGA 499
QY 380 AATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACAGGCAATTCAGCAACAGAGCT 439
DB 498 AATAGAAAAT--ACCCAGNAGTAAAGTTTACTGG-TACCAGGCAATTCAGCACCAGANT 442
QY 440 CTTTCAGAACTGAGGAGAGGTGGAAATCTCAGATCAAGCAGCTGAGGAGAGAGGTG 499
DB 441 CTTTCAGAACTGA-GGAGAGGTGGAAATCTCAGATCAAGCAGCTGAGGAGAGAGGTG 383
QY 500 ATAGAGTAGAAGAGTGGAAAGGCAAAAGAGAAATGAAAAGCAGGCTCAAAACGGA 559
DB 382 ATAGAGTAGAAGAGTGGAAAGGCAAAAGAGAAATGAAAAGCAGGCTCAAAACGGA 323
QY 560 AAAAGTCATATCTTCAAGAAAATCTCTTAAACAGTCCCGGAAATCTCCAGGAGATGAAG 619
DB 322 AAAAGTCATATCTTCAAGAAAATCTCTTAAACAGTCCCGGAAATCTCCAGGAGATGAAG 263
QY 620 ATGACAAAGACTGCAAGAGAGGAAACAAAGAGCAGCTCTGAGGCTGGAGATGCGGCA 679
DB 262 ATGACAAAGACTGCAAGAGAGGAAACAAAGAGCAGCTCTGAGGCTGGAGATGCGGCA 203
QY 680 ACGACAAAGAAACAACTTCAGACTTCAGAAAAACCAAGGAGGACCTAACTACCAT 739

Db 202 ACCGACAGAAACAACTTCAGACTTGCAGAAAAACAGTGAAGGACCTAACTACCAT 143
QY 740 AATGATGCTGATATTAGAGAAACCAAGAGGTTATATGTTGGTGTCTAATATT 799
Db 142 AATGATGCTGATATTAGAGAAACCAAGAGGTTATATGTTGGTGTCTAATATT 83
QY 800 CTTGATTTGATATGAACCAACACATAGTCTTTGTTGTCATTCAGACAGACCCAGTTTGT 859
Db 82 CTTGATTTGATATGAACCAACACATAGTCTTTGTTGTCATTCAGACAGACCCAGTTTGT 23
QY 860 ATGTACAT 867
Db 22 ATGTACCT 15

RESULT 5
US-08-760-745-4
; Sequence 4, Application US/08760745
; Patent No. 5972658
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: LUNG GROWTH FACTOR VARIANT
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,745
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0169 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2376 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 598956
US-08-760-745-4

Query Match 18.9%; Score 193.4; DB 2; Length 2376;
Best Local Similarity 62.5%; Pred. No. 7,1e-46;
Matches 302; Conservative 0; Mismatches 181; Indels 0; Gaps 0;
QY 23 GCGCGCGGAGCGCGCAATTCTGCGCGCGCGGGGGCGGCTCCCGGCATCTTC 82
Db 221 GCGCGAGGAGCG 280
QY 83 GCG 142

Db 281 CCCGCGCGGAGCG 340
QY 143 AGTACAAAGCGCGGACCTGCTTTCGCAAGATGAAGGGCTACCGCACTGGCGGCCC 202
Db 341 AGTACAAATGCGGGGACCTGCTTTCGCAAGATGAAGGGCTACCGCACTGGCGGCCC 400
QY 203 GGATTGATGAACCTCCAGAGGGCGCTGTGAAGCCCTCCAGCAAAACAAGTATCTTCTT 262
Db 401 GGAATTGACGAGATGCTTGAGGCTGCGGTGAATCAACAGCCAAACAATACCAAGTCTTTT 460
QY 263 TTTTGGCACCATGAATCTGCAATTTCTAGGTCCCAAGACCTTTTTCATATTAAGGAT 322
Db 461 TTTTGGGACCCACGAGACGCGCAATTCCTGGGCGCCCAAGACCTTCTTCCCTTACGAGGAT 520
QY 323 ACAAGACAAAGTTTGAAGTCAACAAACGGAAGGATTTAAACGAAGATTTGTCGGAAA 382
Db 521 CCAAGGAAAGTTTGGCAAGGCCCAACAGAGAAAGGGTTTACGAGGGGCTGTGGAGA 580
QY 383 TAGAAATAACCCAGGAGTAAAGTTTACTTGGCTACCGGCAATTCAGCAACAGAGCTCTT 442
Db 581 TCGAGAACACCTTACTGTCAAGGCTTCCGGCTATCAGTCTCCAGAAAAAGAGCTGTG 640
QY 443 CAGAACTGAGGAGAGAGTGAATACTGAGATGCAAGAGTGAAGAGAGAGAGTGA 502
Db 641 TGAAGAGCTCAACACGAGCGCGGAAAGCTGACGGGTGACGGTGTAAAGAGGGGAATG 700
QY 503 GAG 505
Db 701 CAG 703

RESULT 6
US-08-760-745-2
; Sequence 2, Application US/08760745
; Patent No. 5972658
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: LUNG GROWTH FACTOR VARIANT
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,745
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0169 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 869 base pairs
; TYPE: nucleic acid

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2004, 10:16:26 ; Search time 395 Seconds

(without alignments)
9137.614 Million cell updates/sec

Title: US-09-787-328B-3

Perfect score: 1024

Sequence: 1 accgcctgcctcccgccgtt.....actatacatgtctaggatc 1024

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1000	97.7	1898	9	US-09-768-826-22
2	985.6	96.2	1919	9	US-09-925-301-145
3	980	95.7	3805	13	US-10-247-671-8
4	828.4	80.9	1664	9	US-09-876-889-68
5	828.4	80.9	1664	9	US-09-876-889-267
6	778	76.0	2920	11	US-09-866-050A-541
7	778	76.0	2920	15	US-10-152-661-541
8	471	46.0	645	10	US-09-879-536-318
9	248.6	24.3	398	10	US-09-960-352-8612
10	239.6	23.4	288	10	US-09-960-352-1025
11	193.4	18.9	2376	9	US-09-938-885A-4
12	193.4	18.9	2376	15	US-10-207-791-3
13	189	18.5	723	15	US-10-207-791-1
14	189	18.5	2133	9	US-09-925-302-68
15	188.6	18.4	869	9	US-09-938-885A-2

16	173	16.9	3316	13	US-10-133-013-166	Sequence 166, App
17	172.6	16.9	800	9	US-09-910-943-585	Sequence 585, App
18	168.8	15.5	480	11	US-09-918-995-32444	Sequence 32444, A
19	159.8	15.6	1297	9	US-09-768-826-24	Sequence 24, Appl
20	159.8	15.6	2031	11	US-09-823-187-23	Sequence 23, Appl
21	159.8	15.6	2272	11	US-09-946-374-307	Sequence 307, App
22	159.8	15.6	2272	12	US-10-011-795A-307	Sequence 307, App
23	159.8	15.6	2272	13	US-10-015-387A-307	Sequence 307, App
24	159.8	15.6	2272	13	US-10-006-130A-307	Sequence 307, App
25	159.8	15.6	2272	13	US-10-199-672-345	Sequence 345, App
26	159.8	15.6	2272	13	US-10-006-172A-307	Sequence 307, App
27	159.8	15.6	2272	13	US-10-187-749-345	Sequence 345, App
28	159.8	15.6	2272	13	US-10-194-457-345	Sequence 345, App
29	159.8	15.6	2272	13	US-10-184-642-345	Sequence 345, App
30	159.8	15.6	2272	13	US-10-196-747-345	Sequence 345, App
31	159.8	15.6	2272	13	US-10-015-392A-307	Sequence 307, App
32	159.8	15.6	2272	13	US-10-017-253A-307	Sequence 307, App
33	159.8	15.6	2272	13	US-10-173-689-345	Sequence 345, App
34	159.8	15.6	2272	13	US-10-173-690-345	Sequence 345, App
35	159.8	15.6	2272	13	US-10-173-691-345	Sequence 345, App
36	159.8	15.6	2272	13	US-10-173-692-345	Sequence 345, App
37	159.8	15.6	2272	13	US-10-173-694-345	Sequence 345, App
38	159.8	15.6	2272	13	US-10-173-698-345	Sequence 345, App
39	159.8	15.6	2272	13	US-10-173-699-345	Sequence 345, App
40	159.8	15.6	2272	13	US-10-173-707-345	Sequence 345, App
41	159.8	15.6	2272	13	US-10-174-569-345	Sequence 345, App
42	159.8	15.6	2272	13	US-10-174-583-345	Sequence 345, App
43	159.8	15.6	2272	13	US-10-174-587-345	Sequence 345, App
44	159.8	15.6	2272	13	US-10-174-589-345	Sequence 345, App
45	159.8	15.6	2272	13	US-10-174-591-345	Sequence 345, App

ALIGNMENTS

RESULT 1

US-09-768-826-22
; Sequence 22, Application US/09768826
; Patent No. US20020012966A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: 18 human secreted proteins
; FILE REFERENCE: PFS12P1
; CURRENT APPLICATION NUMBER: US/09/768,826
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/US00/22350
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/148,759
; PRIOR FILING DATE: 1999-08-16
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-768-826-22

Query Match	97.7%	Score	1000	DB	9	Length	1898	
Best Local Similarity	99.8%	Pred. No.	7e-267					
Matches 1022	Conservative	0	Mismatches	0	Indels	2	Gaps	2
QY	2	CGCTCGTCCGCCCGCTTTAGCGCGCGGGAGCGC-GCGCAATTCGTCCGCCCGCGGG	60					
Db	30	CGCTCGTCCGCCCGCTTTAGCGCGCGGGAGCGCGCAATTCGTCCGCCCGCGGG	89					
QY	61	GGGCGCGCTCCCGGCATCTTCGCGCGACCAAGGACTACAGGAAGGGAGCGGCTGGG	120					
Db	90	GGGCGCGCTCCCGGCATCTTCGCGCGACCAAGGACTACAGGAAGGGAGCGGCTGGG	149					
QY	121	ATGGCGCGTCCGCGCGCGCGAGTACAAAGCGGGCGACCTGGTCTTCGCCCAAGATGAAG	180					
Db	150	ATGGCGCGTCCGCGCGCGCGAGTACAAAGCGGGCGACCTGGTCTTCGCCCAAGATGAAG	209					

Db 733 GAAAGGACCTAACTACCATAATGCTGCATATTAAGAGAAACCAAGAGGTTATA 792
Qy 781 TGTTGGTGTCTAATAATCTTGATTTGATATGAACCAACACATAGTCTCTGTTGTCAT 840
Db 793 TGTTGGTGTCTAATAATCTTGATTTGATATGAACCAACATAGTCTCTGTTGTCAT 852
Qy 841 TGACAGAACCCAGTTTGTATGTACATATTTCAATATTCCTCTCTGTTGTTGTCGGGGG 900
Db 853 TGACAGAACCCAGTTTGTATGTACATATTTCAATATTCCTCTCTGTTGTTGTCGGGGG 912
Qy 901 AAAAGACATTTAGCCCTTTTAAAGTTTACTGATTTAAATTCATGTTATTTGGTTGTCAT 960
Db 913 AAAAGACATTTAGCCCTTTTAAAGTTTACTGATTTAAATTCATGTTATTTGGTTGTCAT 972
Qy 961 GAAAGTTCCTTAACCACTAAGGATTTATCAAGATTTTGGCGACCTTATACATGTCCTAG 1020
Db 973 GAAAGTTCCTTAACCACTAAGGATTTATCAAGATTTTGGCGACCTTATACATGTCCTAG 1032
Qy 1021 GATC 1024
Db 1033 GATC 1036

RESULT 3

US-10-247-671-8
; Sequence 8, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 8
; LENGTH: 3805
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 2676869CB1
; NAME/KEY: unsure
; LOCATION: 3788
; OTHER INFORMATION: a, t, c, g, or other
US-10-247-671-8

Query Match 95.7%; Score 980; DB 13; Length 3805;
Best Local Similarity 99.1%; Pred. No. 4e-261;
Matches 1017; Conservative 0; Mismatches 5; Indels 4; Gaps 3;
Qy 1 ACCGCTCGTCGCGCGCTTAGCGCCCGCGGGAGCGCGGCAATTCGTGCGCCCGCGGG 60
Db 4 ACCGCTCGTCG-CGCGCTTGAGCGCGCGGGAG-CGCGGCAATTCGTGCGCCCGCGGG 61
Qy 61 GGGCGCGCC--TCCCGCATCTTCGCGCGACCAAGGACTACCAGGAAGGGAGCGGCTG 118
Db 62 GGGCGCGCATTCGCGCAGTCTTCGCGCGACCAAGGACTACCAGGAAGGGAGCGGCTG 121
Qy 119 GGATGCGCGCTCGCGCGCCCGCGAGTACAAAGCGCGGACCTCGTCTTCGCCAAGATGA 178
Db 122 GGATGCGCGCTCGCGCGCCCGCGAGTACAAAGCGCGGACCTCGTCTTCGCCAAGATGA 181
Qy 179 AGGCTACCGGACATGCGCGCGCGCGGATGATGAATCCAGAGGGCGCTGTGAAGCCTC 238
Db 182 AGGCTACCGGACATGCGCGCGCGCGGATGATGAATCCAGAGGGCGCTGTGAAGCCTC 241

Qy 239 CAGCAACAAAGTATCCTATCTCTTTTGGCACCCATGAAATGCAATTTCTAGGTCCCA 298
Db 242 CAGCAACAAAGTATCCTATCTCTTTTGGCACCCATGAAATGCAATTTCTAGGTCCCA 301
Qy 299 AAGACCTTTTCCATATTAAGGAGTACAAGACAAGTTTGGAAAGTCAAAACAAACGGAAG 358
Db 302 AAGACCTTTTCCATATTAAGGAGTACAAGACAAGTTTGGAAAGTCAAAACAAACGGAAG 361
Qy 359 GATTTAAACGAAGGATTTGCGGAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTACC 418
Db 362 GATTTAAACGAAGGATTTGCGGAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTACC 421
Qy 419 AGGCAATTCAGCAACAGAGCTCTTTCAGAAACTCAGGGAGAGGTGGAATACTGCGAGATG 478
Db 422 AGGCAATTCAGCAACAGAGCTCTTTCAGAAACTCAGGGAGAGGTGGAATACTGCGAGATG 481
Qy 479 CAAGCAGTAGGGAAGAGGTGATAGAGTAGAAGATGGAAGAGGCAAGAAAGAAATG 538
Db 482 CAAGCAGTAGGGAAGAGGTGATAGAGTAGAAGATGGAAGAGGCAAGAAAGAAATG 541
Qy 539 AAAAAGCAGGCTCAAAACGGAAGAGTCAATATCTCTCAAGAAATCTCTTAAACAGTCCC 598
Db 542 AAAAAGCAGGCTCAAAACGGAAGAGTCAATATCTCTCAAGAAATCTCTTAAACAGTCCC 601
Qy 599 GGAATCTCCAGGAGATGAAGATGACAAAGACTGCAAAAGAGAGGAAACAAAGAGCAGCT 658
Db 602 GGAATCTCCAGGAGATGAAGATGACAAAGACTGCAAAAGAGAGGAAACAAAGAGCAGCT 661
Qy 659 CTGAGGCTGAGATGCGGGCAACGACACAAGAAACAACTTCAGACTTCGAGAAACCA 718
Db 662 CTGAGGCTGAGATGCGGGCAACGACACAAGAAACAACTTCAGACTTCGAGAAACCA 721
Qy 719 GTCAAGGGACCTTAACCTACCATATGATGCTCATATTAAGAGAAACCAAGAGGTTA 778
Db 722 GTCAAGGGACCTTAACCTACCATATGATGCTCATATTAAGAGAAACCAAGAGGTTA 781
Qy 779 TATGTTTGGTGTCTAATATTTCTGGAATTTGATGAACCAACACATAGTCTTGTGTC 838
Db 782 TATGTTTGGTGTCTAATATTTCTGGAATTTGATGAACCAACACATAGTCTTGTGTC 841
Qy 839 ATTGACAGAACCCAGTTTGTATGTACATTTATTCATATTCCTCTCTGTTGTTGCGGG 898
Db 842 ATTGACAGAACCCAGTTTGTATGTACATTTATTCATATTCCTCTCTGTTGTTGCGGG 901
Qy 899 GGAAGAGACATTTAGCCCTTTTAAAGTTACTGATTTAAATTTCAATTTGTTGTTGTC 958
Db 902 GGAAGAGACATTTAGCCCTTTTAAAGTTACTGATTTAAATTTCAATTTGTTGTTGTC 961
Qy 959 ATGAAGTTGCGCTTAACCACTAAGGATTTATCAAGATTTTTCGCGAGACTTATACATGCT 1018
Db 962 ATGAAGTTGCGCTTAACCACTAAGGATTTATCAAGATTTTTCGCGAGACTTATACATGCT 1021
Qy 1019 AGGATC 1024
Db 1022 AGGATC 1027

RESULT 4

US-09-876-889-68
; Sequence 68, Application US/09876889
; Patent No. US20020076715A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C3
; CURRENT APPLICATION NUMBER: US/09/876,889
; CURRENT FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 353

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 68
; LENGTH: 1664
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-876-889-68

Query Match      80.9%; Score 828.4; DB 9; Length 1664;
Best Local Similarity 98.7%; Pred. No. 3.1e-219;
Matches 835; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 179 AGGGGTACCCGCACTGGCCGCGGATGATGAATCCAGAGGGCGCTGTGAAGCCTC 238
Db 17 AGGGCGGCGGAGTGGCGGATGATGAATCCAGAGGGCGCTGTGAAGCCTC 76
QY 239 CAGCAAAACAAGTATCTATCTCTTTTGGCACCCATGAATCTGATTTCTAGTCCCA 298
Db 77 CAGCAAAACAAGTATCTATCTCTTTTGGCACCCATGAATCTGATTTCTAGTCCCA 136
QY 299 AAGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAACAAACGGAAG 358
Db 137 AAGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAACAAACGGAAG 196
QY 359 GATTTAAACGAAGATTTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACC 418
Db 197 GATTTAAACGAAGATTTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACC 256
QY 419 AGGCAATTCCAGCAACAGAGCTCTTCAGAAACTGAGGAGAAAGTGAATATCTGCAGATG 478
Db 257 AGGCAATTCCAGCAACAGAGCTCTTCAGAAACTGAGGAGAAAGTGAATATCTGCAGATG 316
QY 479 CAAGCAGTGAAGAGAGTGTAGAGTAGAAGAGATGGAAGAGGCAAAAGAAAGATG 538
Db 317 CAAGCAGTGAAGAGAGTGTAGAGTAGAAGAGATGGAAGAGGCAAAAGAAAGATG 376
QY 539 AAAAAAGCAGGCTCAAAACGGAAGAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCC 598
Db 377 AAAAAAGCAGGCTCAAAACGGAAGAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCC 436
QY 599 GGAATCTCCAGGAGATGAAGATGACAAAGATGCAAGAGAGAGAAACAAACAGAGCT 658
Db 437 GGAATCTCCAGGAGATGAAGATGACAAAGATGCAAGAGAGAGAAACAAACAGAGCT 496
QY 659 CTGAGGTTGGAGATGCGGCAACGACACAAAGAAACACAACTTCAGACTTCGAGAAACCA 718
Db 497 CTGAGGTTGGAGATGCGGCAACGACACAAAGAAACACAACTTCAGACTTCGAGAAACCA 556
QY 719 GTGAAGGGAACCTTAACCTATTAATGCTGCATATTAAGAGAAACCAACAGAGGTTA 778
Db 557 GTGAAGGGAACCTTAACCTATTAATGCTGCATATTAAGAGAAACCAACAGAGGTTA 616
QY 779 TATGTTGGTGTCTTAATATCTTGGATTGTATGAACCAACACATAGTCTCTTGTGTC 838
Db 617 TATGTTGGTGTCTTAATATCTTGGATTGTATGAACCAACACATAGTCTCTTGTGTC 676
QY 839 ATTGACAGAACCCAGTTTGTATGTACATTTATTCATTTCTCTCTGTTGTTTGGGG 898
Db 677 ATTGACAGAACCCAGTTTGTATGTACATTTATTCATTTCTCTCTGTTGTTTGGGG 736
QY 899 GGAAAGACATTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTT 958
Db 737 GGAAAGACATTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTT 796
QY 959 ATGAAGTTGCCCTTAACCACTAAGATTAATCAAGATTTTGGCAGACTTATACATGCT 1018
Db 797 ATGAAGTTGCCCTTAACCACTAAGATTAATCAAGATTTTGGCAGACTTATACATGCT 856
QY 1019 AGGATC 1024
Db 857 AGGATC 862
```

```
US-09-876-889-267
; Sequence 267, Application US/09876889
; Patent No. US20020076715A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C3
; CURRENT APPLICATION NUMBER: US/09/876.889
; CURRENT FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 353
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 267
; LENGTH: 1664
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-876-889-267
```

```
Query Match      80.9%; Score 828.4; DB 9; Length 1664;
Best Local Similarity 98.7%; Pred. No. 3.1e-219;
Matches 835; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 179 AGGGCTACCCGCACTGGCCGCGGATGATGAATCCAGAGGGCGCTGTGAAGCCTC 238
Db 17 AGGGCGGCGGAGTGGCGGATGATGAATCCAGAGGGCGCTGTGAAGCCTC 76
QY 239 CAGCAAAACAAGTATCTATCTCTTTTGGCACCCATGAATCTGATTTCTAGTCCCA 298
Db 77 CAGCAAAACAAGTATCTATCTCTTTTGGCACCCATGAATCTGATTTCTAGTCCCA 136
QY 299 AAGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAACAAACGGAAG 358
Db 137 AAGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAACAAACGGAAG 196
QY 359 GATTTAAACGAAGATTTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACC 418
Db 197 GATTTAAACGAAGATTTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACC 256
QY 419 AGGCAATTCCAGCAACAGAGCTCTTCAGAAACTGAGGAGAAAGTGAATATCTGCAGATG 478
Db 257 AGGCAATTCCAGCAACAGAGCTCTTCAGAAACTGAGGAGAAAGTGAATATCTGCAGATG 316
QY 479 CAAGCAGTGAAGAGAGTGTAGAGTAGAAGAGATGGAAGAGGCAAAAGAAAGATG 538
Db 317 CAAGCAGTGAAGAGAGTGTAGAGTAGAAGAGATGGAAGAGGCAAAAGAAAGATG 376
QY 539 AAAAAAGCAGGCTCAAAACGGAAGAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCC 598
Db 377 AAAAAAGCAGGCTCAAAACGGAAGAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCC 436
QY 599 GGAATCTCCAGGAGATGAAGATGACAAAGATGCAAGAGAGAGAAACAAACAGAGCT 658
Db 437 GGAATCTCCAGGAGATGAAGATGACAAAGATGCAAGAGAGAGAAACAAACAGAGCT 496
QY 659 CTGAGGTTGGAGATGCGGCAACGACACAAAGAAACACAACTTCAGACTTCGAGAAACCA 718
Db 497 CTGAGGTTGGAGATGCGGCAACGACACAAAGAAACACAACTTCAGACTTCGAGAAACCA 556
QY 719 GTGAAGGGAACCTTAACCTATTAATGCTGCATATTAAGAGAAACCAACAGAGGTTA 778
Db 557 GTGAAGGGAACCTTAACCTATTAATGCTGCATATTAAGAGAAACCAACAGAGGTTA 616
QY 779 TATGTTGGTGTCTTAATATCTTGGATTGTATGAACCAACACATAGTCTCTTGTGTC 838
Db 617 TATGTTGGTGTCTTAATATCTTGGATTGTATGAACCAACACATAGTCTCTTGTGTC 676
QY 839 ATTGACAGAACCCAGTTTGTATGTACATTTATTCATTTCTCTCTGTTGTTTGGGG 898
Db 677 ATTGACAGAACCCAGTTTGTATGTACATTTATTCATTTCTCTCTGTTGTTTGGGG 736
```

Qy 899 GGAAGAAGACATTTTACGCTTTTAAAGATTACTGATTAATTAATTCATGTTATTTGGTTC 958
Db 737 GGAAGAAGACATTTTACGCTTTTAAAGATTACTGATTAATTAATTCATGTTATTTGGTTC 796
Qy 959 ATGAAGTTGGCTTAAACACATTAAGATTTATCAAGATTTTTCGCGAGACTTTATACATGCT 1018
Db 797 ATGAAGTTGGCTTAAACACATTAAGATTTATCAAGATTTTTCGCGAGACTTTATACATGCT 856
Qy 1019 AGGATC 1024
Db 857 AGGATC 862
RESULT 6
US-09-866-050A-541
; Sequence 541, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866.050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 541
; LENGTH: 2920
; TYPE: DNA
; ORGANISM: Rat
US-09-866-050A-541

Query Match 76.0%; Score 778; DB 11; Length 2920;
Best Local Similarity 89.1%; Pred. No. 4.6e-205;
Matches 912; Conservative 0; Mismatches 95; Indels 17; Gaps 6;

Qy 2 CCGCTCGTCCGCGCGCTTGAAGCGCGCGGAGGCG-GCGCAATTCGTGCGCGCGCGG 60
Db 164 CCGCTCGAGCGCGCGCGCTCGAGGCGCGCGGAGCGCGCGCTCGCTCGCGCGGGA 223
Qy 61 GGGCGGCGCTCCGCGCATCTTCGCGCGGACCAAGGACTACCAAGAGGGGAGCGCTGGG 120
Db 224 GGGCGGCGCTCCGCGCATCTTCGCGCGGACCAAGGACGAGCGGCGGAGGAAACCAACGGG 283
Qy 121 ATGGCGCGCTCCGCGCGCGCGGAGTACAAAGCGGCGGAGCTGGTCTTCGCCAAGATGAAG 180
Db 284 ATGGCGCGCTCCGCGCGCGCGGAGTACAAAGCGGCGGAGCTGGTCTTCGCCAAGATGAAG 343
Qy 181 GGCTACCGCGACTCGCGCGCGCGGATGATGAATCCCGAGAGGCGCTGTGAAGCCTCCA 240
Db 344 GGCTACCGCGACTCGCGCGCGCGGATGATGAATCCCGAGAGGCGCGCTGTGAAGCCTCCA 403
Qy 241 GCAACAGATTCCTATCTCTTTTGGCACCCCATGAACTGAACTTCTAGTCCCAAA 300
Db 404 GCAACAGATTCCTATCTCTTTTGGCACCCCATGAACTGAACTTCTAGTCCCTAAA 463
Qy 301 GACCTTTTCCATATAAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACAAACGGAAGGA 360
Db 464 GACCTTTTCCATATAAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACAAACGGAAGGA 523
Qy 361 TTTAACGAAGATTGTGGAAATAGAAATACCAAGAGTAAAGTTTACTGGCTACCAAG 420
Db 524 TTTAATGAAGATTATGGAAATTTGAAATATATCCAGAGTGAATTTACTGGGTACCAAG 583
Qy 421 GCAATTGAGCAACAGAGCTTTCAGAACTGAGGAGAGTGGAATATCTGCAAGTGA 480
Db 584 ACAATTGAGCAACAGAGCTTTCAGAACTGAGGAGAGTGGAATATCTGCAAGTGA 643

Qy 481 AGCAGTGAGGAAGAGGTGATAGAGTAGAAGATGGAAGGCAAGAAAGAAATGAA 540
Db 644 AGCAGTGAGGAAGAGGTGACAGAGT---AGAAGATGGAAGGCAAGAAAGAAATGAA 700
Qy 541 AAAGCAGGCTCAAAACGGAAAGAGTCAATATCTTCAAGAAATCTCTTAACAGTCCCGG 600
Db 701 AAAGGAGGCTCAAAACGGAAAGAGTCTACACTTCAAGAAAGTCTCTTAACAGTCCCGG 760
Qy 601 AAATCTCCAGAGATGAAGATGACAAAGAGCTGCAAGAGAGGAGGAAACAAAGCAGCTCT 660
Db 761 AAATCTCCAGAGAGCAGAGATGACAAAGAGCTGCAAGAGAGGAGGAAACAAAGCAGCTCT 820
Qy 661 GAGCGTGGAGATCGGCGCAACGACACAAAGAAACACAACTTCAGACTTTCAGAAACCCAGT 720
Db 821 GAGCGCGGAGATGCTGGCAATGACACAAAGAAACACGAGCTTCAGACTTTCAGAAAGCCGCT 880
Qy 721 GAAGGAGCTTAACCTACCAATGAATGAAATGCTGCATATTAAGAGAAACCAAGAAAGTTATA 780
Db 881 GAAGGAGCTTAACCTAATGAAATGCTGCATATTAAGAGAAACCAAGAAAGTTATAAC 940
Qy 781 TGTTTGGTCTCTAAATATCTTGGATTTGATATGAACCAACACATAGTCTCTTGTGTGTCAT 840
Db 941 TGTTTGGTCTCT---GATCTTGGATTTGATATGAACCAAC---AGTCTTGTGTGTCAC 994
Qy 841 TGACAGAACCCAGTTTGTATGTACATATTCATATTCCTCTCTCTGTGTGTGTTTCGGGGGG 900
Db 995 TGACAAAGCCCGAGTGTGTGTGCTATTT---ACATTCCTCTCTCTGTGTGTGTTTCGAAAGGAA 1050
Qy 901 AAAAGACATTTTAGCCCTTTTAAAAGTACTGATTAATTCATGTTATTTGGTTCGAT 960
Db 1051 AAAAGAGACATTTTAAACCTTTTAAAGGTTATTGACTTAATTTATGTTATTTGGTTCGAT 1110
Qy 961 GAAAGTTGCCCTTAACCACTAAGGATTAATCAAGATTTTGGCGAGACTTATACATGCTCTAG 1020
Db 1111 GAAAGTTGCCCTTAACCACTAAGGATTAATCAAGATTTT---GCAGACTTCTCCGTGCTCTAG 1167
Qy 1021 GATC 1024
Db 1168 GCTC 1171

RESULT 7

US-10-152-661-541
; Sequence 541, Application US/10152661
; Publication No. US20030022835A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c5
; CURRENT APPLICATION NUMBER: US/10/152.661
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 09/866,050
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/221,232
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/206,650
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 09/312,283
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/NZ99/00051
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 09/188,930
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 09/069,726
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 541
; LENGTH: 2920
; TYPE: DNA
; ORGANISM: Rat
US-10-152-661-541

Query Match 76.0%; Score 778; DB 15; Length 2920;
Best Local Similarity 89.1%; Pred. No. 4.6e-205;
Matches 912; Conservative 0; Mismatches 95; Indels 17; Gaps 6;

```
QY 2 CGCTCGTCGCGCGCGTGGAGCCCGCGGGAGCGC-GGCAATTCGTGCGCGCGCGG 60
Db 164 CGCTCGAGCGCGCGCGTCCGAGCGCGCGGAGCGCGCGTCCGTCGTCCCGCGGA 223
QY 61 GGGCGCGCTCCCGGCATCTTCGCGGACCAAGGACTACCGAGGAGGCGGCGTGG 120
Db 224 GGGCGCGCTCCCGGCATCTTCGCGGACCAAGGACTACCGAGGAGGCGCGTGAAG 283
QY 121 ATGGCGCGTCCGCGCGCGCGAGTACAAAGCGGGCGACTGTGCTTCGCGCAAGATGAAG 180
Db 284 ATGGCGCGTCCGCGCGCGCGAGTACAAAGCGGGAGACTGTGCTTCGCGCAAGATGAAG 343
QY 181 GGTACCGGCATCGGCGCGCGCGAGTATGATGAATCCAGAGGCGCTGTGAAGCTCCA 240
Db 344 GGTACCGGCATCGGCGCGCGCGAGTATGATGAATCCAGAGGCGCGTGAAGCTCCA 403
QY 241 GCAACAAGTATCTATCTCTTTTGGCACCCATGAACTGCATTTCTAGTCCCAA 300
Db 404 GCAACAAGTATCTATCTCTTTTGGCACCCATGAACTGCATTTCTAGTCCCAA 463
QY 301 GACTTTTTCATATAGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAAGGAAAGGA 360
Db 464 GACTTTTTCATATAGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAAGGAAAGGA 523
QY 361 TTTAAAGAGATTTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGCTACCAAG 420
Db 524 TTTAAAGAGATTTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGCTACCAAG 583
QY 421 GCAATTCAGACAGAGCTCTTCAGAAACTGAGGAGAGGTCGAAATCTCAGATGCA 480
Db 584 GCAATTCAGACAGAGCTCTTCAGAAACTGAGGAGAGGTCGAAATCTCAGATGCA 643
QY 481 AGCAGTGAGGAGGAGGATGATGAGTGAAGAGATGGAAGAGGCAAAAGGAAAGATGAA 540
Db 644 AGCAGTGAGGAGGAGGATGATGAGTGAAGAGATGGAAGAGGCAAAAGGAAAGATGAA 700
QY 541 AAAGAGGCTCAAAACGGAAGATCATATATCTTCAAGAAATCTCTTAAACAGTCCCG 600
Db 701 AAAGAGGCTCAAAACGGAAGATCATATATCTTCAAGAAATCTCTTAAACAGTCCCG 760
QY 601 AATCTCCAGGAGATCAAGATGACAAAGACTGCAAGAGAGGAAACAAAGAGCAGCTCT 660
Db 761 AATCTCCAGGAGATCAAGATGACAAAGACTGCAAGAGAGGAAACAAAGAGCAGCTCT 820
QY 661 GAGGAGGAGATCGCGGCAACGACACAAAGAACTTTCAGACTTTCAGAGAAACCAAGT 720
Db 821 GAGGAGGAGATCGCGGCAACGACACAAAGAACTTTCAGACTTTCAGAGAAACCAAGT 780
QY 721 GAAGGAGCTTAATCAATGAATGCTGCAATTTAAGAGAAACCAAGAGAGGTTATA 780
Db 881 GAAGGAGCTTAATCAATGAATGCTGCAATTTAAGAGAAACCAAGAGAGGTTATA 940
QY 781 TGTGTTGTTGTTCTAATATCTTGGATTCATATGACCAACATAGTCTTGTGTCAT 840
Db 941 TGTGTTGTTGTTCTAATATCTTGGATTCATATGACCAACATAGTCTTGTGTCAT 994
QY 841 TGACAGAAACCCAGTTTGTATGTACATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Db 995 TGACAGAAACCCAGTTTGTGTTGTCATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1050
QY 901 AAAAGACATTTAGCCCTTTTAAAGTTACTGATTTAATTTCTGATTTTGTGTCAT 960
Db 1051 AAAAGACATTTTAAACCTTTTAAAGTTATTGACTTAATTTTATGTTTATTTGTTGTCAT 1110
```

```
QY 961 GAAGTTGCGCTTAACCACTAAGGATTATCAAGATTTTTCGCGAGACTTATACATGCTTAG 1020
Db 1111 GAAGTTGCGCTTAACCACTAAGGATTATCAAGATTTTTCGCGAGACTTATACATGCTTAG 1167
QY 1021 GATC 1024
Db 1168 GCTC 1171
```

RESULT 8

US-09-879-536-318/c
; Sequence 318, Application US/09879536
; Patent No. US20020144298A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/879,536
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/088,801
; PRIOR FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 318
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(645)
; OTHER INFORMATION: n = A,T,C or G
US-09-879-536-318

Query Match 46.0%; Score 471; DB 10; Length 645;
Best Local Similarity 95.8%; Pred. No. 3.4e-120;
Matches 525; Conservative 0; Mismatches 18; Indels 5; Gaps 4;

```
QY 320 AGTACAAAGACAGAGTTTCGAAAGTCAAAACAAACGGAAGGATTTTAAACGAAGGATTGTGG 379
Db 557 AGGTCCAANGCCAGTTTGAAAAGTC-AAACCAACGGAAGGATTTTAAACGAAGGATTGTGA 499
QY 380 AAATAGAAATAAACCCAGGAGTAAAGTTTACTGGTACAGGCAATTCAGCAACAGAGCT 439
Db 498 ATAGAAAT--ACCCAGNAGTAAAGTTTACTGG--TACCAGGCATTCAGCACAGAGNT 442
QY 440 CTTTCAAGAACTGAGGAGGAGAGTGGAAATATCTGCAAGATGCAAGAGTGAAGAGGTTG 499
Db 441 CTTTCAAGAACTGA--GGAGAGGTTGGAAATATCTGCAAGATGCAAGAGTGAAGAGGTTG 383
QY 500 ATAGAGTAGAAGAGATGGAAGGCAAAAGCAAAAGCAAAAGCAAAAGCAAGCTCAAAACCGA 559
Db 382 ATAGAGTAGAAGAGATGGAAGGCAAAAGCAAAAGCAAAAGCAAAAGCAAGCTCAAAACCGA 323
QY 560 AAAAGCTATATATCTTCAAGAAATCTCTTAAACAGTCCCGGAAATCTCCAGGAGATGAAG 619
Db 322 AAAAGCTATATATCTTCAAGAAATCTCTTAAACAGTCCCGGAAATCTCCAGGAGATGAAG 263
QY 620 ATGACAAAGACTGCAAAAGAGAGGAAACAAAGCAAGCTCTGAGGGTGAAGATGCGGCA 679
Db 262 ATGACAAAGACTGCAAAAGAGAGGAAACAAAGCAAGCTCTGAGGGTGAAGATGCGGCA 203
```


5 1013 98.9 1830 50 US-10-342-887-1603 Sequence 1603, Ap
6 1013 98.9 1998 26 US-09-606-776-5006 Sequence 5006, Ap
7 1013 98.9 2342 28 US-09-644-869-3226 Sequence 9226, Ap
8 1013 98.9 2342 28 US-09-644-870-9669 Sequence 9669, Ap
9 1013 98.9 2342 28 US-09-649-162-9840 Sequence 9840, Ap
10 1013 98.9 2342 28 US-09-652-109-9938 Sequence 9938, Ap
11 1013 98.9 2342 28 US-09-652-816-9425 Sequence 9425, Ap
12 1013 98.9 2342 30 US-09-698-013-6892 Sequence 6892, Ap
13 1013 98.9 2342 32 US-09-726-175-3077 Sequence 3077, Ap
14 1013 98.9 2342 32 US-09-726-806-5292 Sequence 5292, Ap
15 1012 98.8 1909 33 US-09-770-173-2543 Sequence 2543, Ap
16 1012 98.8 1973 27 US-09-629-469A-11306 Sequence 11306, A
17 1012 98.8 2669 20 US-09-371-168-7090 Sequence 7090, Ap
18 1012 98.8 2669 28 US-09-644-867-5952 Sequence 5952, Ap
19 1012 98.8 2669 28 US-09-644-868-7420 Sequence 7420, Ap
20 1012 98.8 2669 28 US-09-644-869-7588 Sequence 7588, Ap
21 1012 98.8 2669 28 US-09-644-870-7347 Sequence 7347, Ap
22 1012 98.8 2669 28 US-09-644-871-7477 Sequence 7477, Ap
23 1012 98.8 2669 28 US-09-649-162-8302 Sequence 8302, Ap
24 1012 98.8 2669 28 US-09-649-164-7849 Sequence 7849, Ap
25 1012 98.8 2669 28 US-09-652-109-8037 Sequence 8037, Ap
26 1012 98.8 2669 28 US-09-652-126-8504 Sequence 8504, Ap
27 1012 98.8 2669 28 US-09-652-814-9037 Sequence 9037, Ap
28 1012 98.8 2669 28 US-09-652-913-8643 Sequence 8643, Ap
29 1012 98.8 2669 28 US-09-652-916-8405 Sequence 8405, Ap
30 1012 98.8 2669 28 US-09-652-914-7662 Sequence 7662, Ap
31 1012 98.8 2669 34 US-09-801-833-7090 Sequence 7090, Ap
32 1012 98.8 3193 20 US-09-396-885-4251 Sequence 4251, Ap
33 1012 98.8 3193 20 US-09-396-885A-4251 Sequence 4251, Ap
34 1012 98.8 3193 20 US-09-399-932-4843 Sequence 4843, Ap
35 1000 97.7 1898 1 PCT-US00-22350-22 Sequence 22, Appl
36 1000 97.7 1898 2 PCT-US00-22350-22 Sequence 22, Appl
37 1000 97.7 1898 33 US-09-768-826-22 Sequence 22, Appl
38 985.6 96.2 1919 1 PCT-US00-05882-145 Sequence 145, Appl
39 985.6 96.2 1919 2 PCT-US00-05882-145 Sequence 145, Appl
40 985.6 96.2 1919 39 US-09-925-301-145 Sequence 145, Appl
41 980 95.7 3805 15 US-09-925-412-4 Sequence 4, Appl
42 980 95.7 3805 48 US-10-247-671-8 Sequence 8, Appl
43 980 95.7 3805 87 US-60-323-784-8 Sequence 8, Appl
44 980 95.7 4711 82 US-60-278-232-5507 Sequence 6507, Ap
45 980 95.7 4711 87 US-60-324-185-21767 Sequence 21767, A

ALIGNMENTS

RESULT 1
US-09-787-328A-3
; Sequence 3, Application US/09787328A
; GENERAL INFORMATION:
; APPLICANT: Yu, Long
; APPLICANT: Zhang, Honglai
; APPLICANT: Fu, Qiang
; APPLICANT: Zhao, Yong
; APPLICANT: Tu, Qiang
; TITLE OF INVENTION: NEW HUMAN HEPATOMA-DERIVED GROWTH FACTOR ENCODING SEQUENCE AND TH
; FILE REFERENCE: 9548.50USWO
; CURRENT APPLICATION NUMBER: US/09/787,328A
; CURRENT FILING DATE: 2001-08-23
; PRIOR APPLICATION NUMBER: PCT/CN99/00139
; PRIOR FILING DATE: 1999-09-06
; PRIOR APPLICATION NUMBER: CN 98119758.2
; PRIOR FILING DATE: 1998-09-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1024
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-787-328A-3

Query Match 100.0%; Score 1024; DB 33; Length 1024;

Best Local Similarity 100.0%; Pred. No. 3.6e-223;
Matches 1024; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCGCTGTCGCGCGCGCTTGAAGCCCGCGGGAGCGCGCAATTCTGTCGGCCGCGGG 60
Db 1 ACCGCTGTCGCGCGCGCTTGAAGCCCGCGGGAGCGCGCAATTCTGTCGGCCGCGGG 60
QY 61 GGGGCGGCTCCCGGCACTCTTCGCGGGACCAAGGACTACCAGGAAGGGAGCGGCTGG 120
Db 61 GGGGCGGCTCCCGGCACTCTTCGCGGGACCAAGGACTACCAGGAAGGGAGCGGCTGG 120
QY 121 ATGCGCGCTCGCGCGCGCGAGTACAAAGCGGGCGACCTGCTGCTTCGCCAAGATGAAG 180
Db 121 ATGCGCGCTCGCGCGCGCGAGTACAAAGCGGGCGACCTGCTGCTTCGCCAAGATGAAG 180
QY 181 GGCTACCCGCACTCGCGCGCGCGGATGTAATGAACTCCAGAGGGCGCTGTGAAGCTCCA 240
Db 181 GGCTACCCGCACTCGCGCGCGCGGATGTAATGAACTCCAGAGGGCGCTGTGAAGCTCCA 240
QY 241 GCAACAGTATCTATCTCTCTTTTGGCACCCATGAACTGCAATTTCTAGGTCCCAAA 300
Db 241 GCAACAGTATCTATCTCTTTTGGCACCCATGAACTGCAATTTCTAGGTCCCAAA 300
QY 301 GACCTTTTTCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAGGA 360
Db 301 GACCTTTTTCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAGGA 360
QY 361 TTTAACAGAGGATTTGGGAAATAGAAATAACCCAGGAGTAAAGTTTACTTGGCTACCAG 420
Db 361 TTTAACAGAGGATTTGGGAAATAGAAATAACCCAGGAGTAAAGTTTACTTGGCTACCAG 420
QY 421 GCAATTCAGCAACAGAGCTCTTCAGAACTCAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
Db 421 GCAATTCAGCAACAGAGCTCTTCAGAACTCAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
QY 481 AGCAGTGAAGAGAGGAGTATAGAGTGAAGAGAGTGAAGAGGAGGAGGAGGAGGAGGAG 540
Db 481 AGCAGTGAAGAGAGGAGTATAGAGTGAAGAGAGTGAAGAGGAGGAGGAGGAGGAGGAG 540
QY 541 AAAGAGGCTCAAAACGGAAGAGTATATCTTCAAGAAATCTCTTAAACAGTCCCGG 600
Db 541 AAAGAGGCTCAAAACGGAAGAGTATATCTTCAAGAAATCTCTTAAACAGTCCCGG 600
QY 601 AAATCTCCAGGAGATGAAGTACAAAGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGG 660
Db 601 AAATCTCCAGGAGATGAAGTACAAAGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGG 660
QY 661 GAGGCTGGAGATCGGGCAACGACACAAAGAACCAACTTCAGACTTGCAGAAACCAAGT 720
Db 661 GAGGCTGGAGATCGGGCAACGACACAAAGAACCAACTTCAGACTTGCAGAAACCAAGT 720
QY 721 GAAGGACCTAACTACCAATATGAATGCTGATTAAGAGAAACCAAGAGAGGTTATA 780
Db 721 GAAGGACCTAACTACCAATATGAATGCTGATTAAGAGAAACCAAGAGAGGTTATA 780
QY 781 TGTGTTGTTGCTAATATCTTGGATTTGATGATCAACCAACACATAGTCTTGTGTCAT 840
Db 781 TGTGTTGTTGCTAATATCTTGGATTTGATGATCAACCAACACATAGTCTTGTGTCAT 840
QY 841 TGACAGAACCCAGTTGATGATCAATTTTCAATATCTCTCTGTTGTTTTCGGGGGG 900
Db 841 TGACAGAACCCAGTTGATGATCAATTTTCAATATCTCTCTGTTGTTTTCGGGGGG 900
QY 901 AAAAGACATTTAGCTTTTAAAGTTACTGATTAATTTCAATTTTCAATTTTGGTTGAT 960
Db 901 AAAAGACATTTAGCTTTTAAAGTTACTGATTAATTTCAATTTTCAATTTTGGTTGAT 960
QY 961 GAAAGTTCCTTAACCAAGGATTAACAGATTTTGGCAGACTTATACATGCTCTAG 1020
Db 961 GAAAGTTCCTTAACCAAGGATTAACAGATTTTGGCAGACTTATACATGCTCTAG 1020
QY 1021 GATC 1024
|||||

Db	1021	GATC 1024	
RESULT 2			
US-09-787-328b-3			
; Sequence 3, Application US/09787328B			
; GENERAL INFORMATION:			
; APPLICANT: Yu, Long			
; APPLICANT: Zhang, Honglai			
; APPLICANT: Fu, Qiang			
; APPLICANT: Zhao, Yong			
; APPLICANT: Tu, Qiang			
; TITLE OF INVENTION: NEW HUMAN HEPATOMA-DERIVED GROWTH FACTOR ENCODING SEQUENCE AND			
; TITLE OF INVENTION: POLYPEPTIDE ENCODED BY SUCH DNA SEQUENCE AND PRODUCING METHOD TH			
; FILE REFERENCE: 9548.50USWO			
; CURRENT APPLICATION NUMBER: US/09/787,328B			
; PRIOR FILING DATE: 2001-03-16			
; PRIOR APPLICATION NUMBER: PCT/CN99/00139			
; PRIOR FILING DATE: 1999-09-06			
; PRIOR FILING DATE: 1998-09-22			
; NUMBER OF SEQ ID NOS: 10			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 3			
; LENGTH: 1024			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-787-328b-3			
Query Match 100.0%; Score 1024; DB 33; Length 1024;			
Best Local Similarity 100.0%; Pred. No. 3.6e-223;			
Matches 1024; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ACCGCTCGTCCGCCGCGTTCAGGCGCGCGGAGCGCGCAATTCGTCGCGCGCGG 60	
Db	1		
Qy	61	GGGGGGCGCTCCCGGCATCTTCGCGCGACCAAGGACTACCAAGGAGGCGGCGTGG 120	
Db	61	GGGGGGCGCTCCCGGCATCTTCGCGCGACCAAGGACTACCAAGGAGGCGGCGTGG 120	
Qy	121	ATGGCGCGTCCGGCGCGCGGATCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180	
Db	121	ATGGCGCGTCCGGCGCGCGGATCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180	
Qy	181	GGTACCGCGCTCGCGCGCGGCGGATTCGATGATGATGATGATGATGATGATGATG 240	
Db	181	GGTACCGCGCTCGCGCGCGGCGGATTCGATGATGATGATGATGATGATGATGATG 240	
Qy	241	GCAACAAGTATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300	
Db	241	GCAACAAGTATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300	
Qy	301	GACCTTTTCCATATAAGGAGTACAAAGCAAGTTGGAAAGTCAAAACCGAAGGA 360	
Db	301	GACCTTTTCCATATAAGGAGTACAAAGCAAGTTGGAAAGTCAAAACCGAAGGA 360	
Qy	361	TTTAACGAAGGATGTGGGAAATAGAAATACCCAGGAGTAAAGTTTACTGCTACC 420	
Db	361	TTTAACGAAGGATGTGGGAAATAGAAATACCCAGGAGTAAAGTTTACTGCTACC 420	
Qy	421	GCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGAGAGAGGTTGAAATCTCAGAT 480	
Db	421	GCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGAGAGAGGTTGAAATCTCAGAT 480	
Qy	481	AGCAGTGAAGAAAGGTTGATAGATGAGTGAAGAGATGGAAGGCAAAAGAAAGTGA 540	
Db	481	AGCAGTGAAGAAAGGTTGATAGATGAGTGAAGAGATGGAAGGCAAAAGAAAGTGA 540	
Qy	541	AAAGCAGGCTCAAAACGGAAGGATCATATCTTCAAGAAATCTCTTAAACAGTCCCG 600	
Db	541	AAAGCAGGCTCAAAACGGAAGGATCATATCTTCAAGAAATCTCTTAAACAGTCCCG 600	
RESULT 3			
PCT-US02-18947-1603			
; Sequence 1603, Application PC/TUS0218947			
; GENERAL INFORMATION:			
; APPLICANT: Rosetta Inpharmatics			
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients			
; FILE REFERENCE: 9301-175-228			
; CURRENT APPLICATION NUMBER: PCT/US02/18947			
; PRIOR FILING DATE: 2002-06-14			
; PRIOR APPLICATION NUMBER: 60/380,770			
; PRIOR FILING DATE: 2002-05-14			
; NUMBER OF SEQ ID NOS: 2699			
; SEQ ID NO 1603			
; LENGTH: 1830			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; PUBLICATION INFORMATION:			
; DATABASE ACCESSION NUMBER: NM 016073			
; DATABASE ENTRY DATE: 2001-06-18			
PCT-US02-18947-1603			
Query Match 98.9%; Score 1013; DB 1; Length 1830;			
Best Local Similarity 99.9%; Pred. No. 1.4e-220;			
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;			
Qy	1	ACCGCTCGTCCGCCGCGCTTCAGGCGCGCGGAGCGCGCAATTCGTCGCGCGCGG 59	
Db	4	ACCGCTCGTCCGCCGCGCTTCAGGCGCGCGGAGCGCGCAATTCGTCGCGCGCGG 63	
Qy	60	GGGGGGCGCTCCCGGCATCTTCGCGCGACCAAGGACTACCAAGGAGGAGCGGCTGG 119	
Db	64	GGGGGGCGCTCCCGGCATCTTCGCGCGACCAAGGACTACCAAGGAGGAGCGGCTGG 123	
Qy	120	GATGCGCGCTCCCGCGCGCGCGAGTACAAAGCGCGCGGCGGCTTCGCGCAAGATGAA 179	
Db	124	GATGCGCGCTCCCGCGCGCGCGAGTACAAAGCGCGCGGCGGCTTCGCGCAAGATGAA 183	
Qy	180	GGGCTACCGGCACTGGCGCGCGCGGATTCGATGAACTCCAGAGGCGGCTGTGAAGCTCC 239	
Db	184	GGGCTACCGGCACTGGCGCGCGCGGATTCGATGAACTCCAGAGGCGGCTGTGAAGCTCC 243	


```
Qy 240 AGCAAAAGATATCTATCTCTTTTGGCACCCATGAATCTGATTTCTAGTCCCAA 299
    |||||
Db 244 AGCAAAAGATATCTATCTCTTTTGGCACCCATGAATCTGATTTCTAGTCCCAA 303
    |||||
Qy 300 AGACTTTTCCATATAGAGATCAAGACAAAGTTGGAAAGTCAACAAACGGAAGG 359
    |||||
Db 304 AGACTTTTCCATATAGAGATCAAGACAAAGTTGGAAAGTCAACAAACGGAAGG 363
    |||||
Qy 360 ATTTAAAGAAAGTTGGGAAATAGAAAATACCCAGGATGAAGTTTACTGCTACCA 419
    |||||
Db 364 ATTTAAAGAAAGTTGGGAAATAGAAAATACCCAGGATGAAGTTTACTGCTACCA 423
    |||||
Qy 420 GGCAATTCAGCAACAGAGCTCTTCAGAAATCGAGGAGAAAGTCTGAGATGC 479
    |||||
Db 424 GGCAATTCAGCAACAGAGCTCTTCAGAAATCGAGGAGAAAGTCTGAGATGC 483
    |||||
Qy 480 AAGCAGTGAGAAAGATGATAGATAGAAAGATGGAAGGCAAGAAAGAAATGA 539
    |||||
Db 484 AAGCAGTGAGAAAGATGATAGATAGAAAGATGGAAGGCAAGAAAGAAATGA 543
    |||||
Qy 540 AAAAGCAGGCTCAAAACGGAAGATGATAGATGATGATGATGATGATGATGATG 599
    |||||
Db 544 AAAAGCAGGCTCAAAACGGAAGATGATAGATGATGATGATGATGATGATGATG 603
    |||||
Qy 600 GAAATCTCCAGAGATGAAGATGATGATGATGATGATGATGATGATGATGATG 659
    |||||
Db 604 GAAATCTCCAGAGATGAAGATGATGATGATGATGATGATGATGATGATGATG 663
    |||||
Qy 660 TGAGGGTGAGATGCGGGCAACGACACAAAGAAACACAACTTCAGACTTCAGAT 719
    |||||
Db 664 TGAGGGTGAGATGCGGGCAACGACACAAAGAAACACAACTTCAGACTTCAGAT 723
    |||||
Qy 720 TGAAGGGACCTAACTACCAATAATGAATGATGATGATGATGATGATGATGATG 779
    |||||
Db 724 TGAAGGGACCTAACTACCAATAATGAATGATGATGATGATGATGATGATGATG 783
    |||||
Qy 780 ATGTTTGGTTGCTTAATATTTCTGGATTTGATGATGATGATGATGATGATGAT 839
    |||||
Db 784 ATGTTTGGTTGCTTAATATTTCTGGATTTGATGATGATGATGATGATGATGAT 843
    |||||
Qy 840 TTGACAGAACCCAGTTTGTATGATGATGATGATGATGATGATGATGATGATG 899
    |||||
Db 844 TTGACAGAACCCAGTTTGTATGATGATGATGATGATGATGATGATGATGATG 903
    |||||
Qy 900 GAAAGACATTTTACCTTTTAAAGTTTACTGATTTTAAATTTTAAATTTTAAAT 959
    |||||
Db 904 GAAAGACATTTTACCTTTTAAAGTTTACTGATTTTAAATTTTAAATTTTAAAT 963
    |||||
Qy 960 TGAAGTTGCCCTTAACCACTAAGGATTTTAAAGTTTACTGATTTTAAATTTTAA 1019
    |||||
Db 964 TGAAGTTGCCCTTAACCACTAAGGATTTTAAAGTTTACTGATTTTAAATTTTAA 1023
    |||||
Qy 1020 GGATC 1024
    |||||
Db 1024 GGATC 1028
    |||||
```

RESULT 4

```
US-10-172-118-1603
; Sequence 1603, Application US/10172118
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
```

```
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1603
; LENGTH: 1830
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_016073
; DATABASE ENTRY DATE: 2001-06-18
; US-10-172-118-1603
```

```
Query Match 98.9%; Score 1013; DB 47; Length 1830;
Best Local Similarity 99.9%; Pred. No. 1.4e-220;
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
Qy 1 ACCGCTGTCGCGCCCGCTTGAAGCCGCGCGGAGGCGCG-CGCAATTCGTGCGCCCGG 59
Db 4 ACCGCTGTCGCGCCCGCTTGAAGCCGCGCGGAGGCGCGCAATTCGTGCGCCCGG 63
    |||||
Qy 60 GGGGCGGCGCTCCCGGCGATCTTCGCGCGACCAAGNACTACCAAGGAGGAGCGGCTGG 119
    |||||
Db 64 GGGGCGGCGCTCCCGGCGATCTTCGCGCGACCAAGNACTACCAAGGAGGAGCGGCTGG 123
    |||||
Qy 120 GATGGCGCGTCCGCGGCGCGCGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 179
    |||||
Db 124 GATGGCGCGTCCGCGGCGCGCGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 183
    |||||
Qy 180 GGGCTACCGCACTGCGCGCGCGCGGATTTGATGAACTCCCAAGAGGCGCTGTGAAGCTCC 239
    |||||
Db 184 GGGCTACCGCACTGCGCGCGCGCGGATTTGATGAACTCCCAAGAGGCGCTGTGAAGCTCC 243
    |||||
Qy 240 AGCAAAACAGTATCTCTATCTTTTGGCACCCATGAATCTGATTTCTAGTCCCAA 299
    |||||
Db 244 AGCAAAACAGTATCTCTATCTTTTGGCACCCATGAATCTGATTTCTAGTCCCAA 303
    |||||
Qy 300 AGACTTTTCCATATAGAGATCAAGACAAAGTTGGAAAGTCAACAAACGGAAGG 359
    |||||
Db 304 AGACTTTTCCATATAGAGATCAAGACAAAGTTGGAAAGTCAACAAACGGAAGG 363
    |||||
Qy 360 ATTTAAAGAAAGTTGGGAAATAGAAAATAAAGAGGATGATGATGATGATGATGATG 419
    |||||
Db 364 ATTTAAAGAAAGTTGGGAAATAGAAAATAAAGAGGATGATGATGATGATGATGATG 423
    |||||
Qy 420 GCGAATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTTGGAATCTGAGATGC 479
    |||||
Db 424 GCGAATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTTGGAATCTGAGATGC 483
    |||||
Qy 480 AAGCAGTGAGAAAGATGATAGATGATGATGATGATGATGATGATGATGATGATG 539
    |||||
Db 484 AAGCAGTGAGAAAGATGATAGATGATGATGATGATGATGATGATGATGATGATG 543
    |||||
Qy 540 AAAAGCAGGCTCAAAACGGAAGATGATAGATGATGATGATGATGATGATGATG 599
    |||||
Db 544 AAAAGCAGGCTCAAAACGGAAGATGATAGATGATGATGATGATGATGATGATG 603
    |||||
Qy 600 GAAATCTCCAGGAGATGAAGATGATGATGATGATGATGATGATGATGATGATG 659
    |||||
Db 604 GAAATCTCCAGGAGATGAAGATGATGATGATGATGATGATGATGATGATGATG 663
    |||||
Qy 660 TGAGGGTGAGATGCGGGCAACGACACAAAGAAACACAACTTCAGACTTCAGAT 719
    |||||
Db 664 TGAGGGTGAGATGCGGGCAACGACACAAAGAAACACAACTTCAGACTTCAGAT 723
    |||||
Qy 720 TGAAGGGACCTAACTACCAATAATGAATGATGATGATGATGATGATGATGATG 779
    |||||
Db 724 TGAAGGGACCTAACTACCAATAATGAATGATGATGATGATGATGATGATGATG 783
    |||||
Qy 780 ATGTTTGGTTGCTTAATATTTCTGGATTTGATGATGATGATGATGATGATGAT 839
    |||||
Db 784 ATGTTTGGTTGCTTAATATTTCTGGATTTGATGATGATGATGATGATGATGAT 843
    |||||
```

```
QY 840 TTGACAGAACCCAGTTTGTATGATACATATTCATATTCCTCTCTGTGTGTTTCGGGG 899
Db 844 TTGACAGAACCCAGTTTGTATGATGATACATATTCATATTCCTCTCTGTGTGTTTCGGGG 903
QY 900 GAAAGACATTTAGCTTTTAAAGTTTACTGATTTAATTTTCATGTTTGGTTGCA 959
Db 904 GAAAGACATTTAGCTTTTAAAGTTTACTGATTTAATTTTCATGTTTGGTTGCA 963
QY 960 TGAAGTTGCCCTTAAACCACTAAGGATTTATCAAGATTTTTCGCGAGACTTATACATGCTCA 1019
Db 964 TGAAGTTGCCCTTAAACCACTAAGGATTTATCAAGATTTTTCGCGAGACTTATACATGCTCA 1023
QY 1020 GGATC 1024
Db 1024 GGATC 1028

RESULT 5
US-10-342-887-1603
; Sequence 1603, Application US/10342887
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1603
; LENGTH: 1830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1603

Query Match 98.9%; Score 1013; DB 50; Length 1830;
Best Local Similarity 99.9%; Pred. No. 1.4e-220;
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ACCGCTCGTCCGCGCCGCTTGAGCCCGCGGGAGCGCG-CGCAATTTCTGTCGGCCCGCG 59
Db 4 ACCGCTCGTCCGCGCCGCTTGAGCCCGCGGGAGCGCGCGCAATTTCTGTCGGCCCGCG 63
QY 60 GGGGGCGGCTCCCGGCATCTTCGCGCGACCAAGGATACCAAGAGGGGAGCGCGCTGG 119
Db 64 GGGGGCGGCTCCCGGCATCTTCGCGCGACCAAGGATACCAAGAGGGGAGCGCGCTGG 123
QY 120 GATGGCGCTCCGCGCCCGCGAGTACAAAGCGGGGACCTGCTTCCGCAAGATGAA 179
Db 124 GATGGCGCTCCGCGCCCGCGAGTACAAAGCGGGGACCTGCTTCCGCAAGATGAA 183
QY 180 GGGCTACCCGCTCCGCGCCCGCGATTTGATGAATCCCAAGCGGGGCTGTGAAGCTCC 239
Db 184 GGGCTACCCGCTCCGCGCCCGCGATTTGATGAATCCCAAGCGGGGCTGTGAAGCTCC 243
QY 240 AGCAACAAGATCTCTATCTTTTGGCAACCCATGAACTTCTAGTCCCAA 299
Db 244 AGCAACAAGATCTCTATCTTTTGGCAACCCATGAACTTCTAGTCCCAA 303
QY 300 AGACCTTTTCCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACGGAAG 359
Db 304 AGACCTTTTCCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACGGAAG 363
```

```
QY 360 ATTAAACGAAGGATTGTGGAAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTACCA 419
Db 364 ATTAAACGAAGGATTGTGGAAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTACCA 423
QY 420 GGAATTCAGCAACAGAGCTCTTCAGAAACTGAGGAGAGAGGTGGAAAATCTTCAGATGC 479
Db 424 GGAATTCAGCAACAGAGCTCTTCAGAAACTGAGGAGAGAGGTGGAAAATCTTCAGATGC 483
QY 480 AGCAGTGAAGGAGAGGTGATAGATGAGAGAGATGAGAGAGGCAAGAGCAAGAGAGATGA 539
Db 484 AGCAGTGAAGGAGAGGTGATAGATGAGAGAGATGAGAGAGGCAAGAGAGAGATGA 543
QY 540 AAAAGCAGGCTCAAAAACGGAAAAAGTCAATATCTTCAAAAGAAATCTCTTAAACAGTCCCG 599
Db 544 AAAAGCAGGCTCAAAAACGGAAAAAGTCAATATCTTCAAAAGAAATCTCTTAAACAGTCCCG 603
QY 600 GAAATCTCAGGAGATGAAAGATGAAAGACTGCAAGAGAGAGAGAGAGAGAGAGAGAGCTC 659
Db 604 GAAATCTCAGGAGATGAAAGATGAAAGACTGCAAGAGAGAGAGAGAGAGAGAGAGAGCTC 663
QY 660 TGAGGCTGAGAGATGCGGGCAACGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 719
Db 664 TGAGGCTGAGAGATGCGGGCAACGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 723
QY 720 TGAAGGAGCTAACTAACCATTAATGATGCTGCATATTAAGAGAGAGAGAGAGAGAGAGAG 779
Db 724 TGAAGGAGCTAACTAACCATTAATGATGCTGCATATTAAGAGAGAGAGAGAGAGAGAGAG 783
QY 780 ATGTTTGGTGTCTAATATCTTGGATTTGATATGAACCAACATAGTCTTGTGTTGTC 839
Db 784 ATGTTTGGTGTCTAATATCTTGGATTTGATATGAACCAACATAGTCTTGTGTTGTC 843
QY 840 TTGACAGAACCCAGTTTGTATGATACATATTCATATTCCTCTCTGTGTGTTTCGGGG 899
Db 844 TTGACAGAACCCAGTTTGTATGATACATATTCATATTCCTCTCTGTGTGTTTCGGGG 903
QY 900 GAAAGACATTTTAGCTTTTAAAGTTTACTGATTTAATTTTCATGTTTGGTTGCA 959
Db 904 GAAAGACATTTTAGCTTTTAAAGTTTACTGATTTAATTTTCATGTTTGGTTGCA 963
QY 960 TGAAGTTGCCCTTAAACCACTAAGGATTTATCAAGATTTTTCGCGAGACTTATACATGCTCA 1019
Db 964 TGAAGTTGCCCTTAAACCACTAAGGATTTATCAAGATTTTTCGCGAGACTTATACATGCTCA 1023
QY 1020 GGATC 1024
Db 1024 GGATC 1028

RESULT 6
US-09-606-776-5006
; Sequence 5006, Application US/09606776
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Myers, Paul
; APPLICANT: Gearing, David P.
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1129-001
; CURRENT APPLICATION NUMBER: US/09/606,776
; PRIOR FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/141,578
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/141,379
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: 60/141,138
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: 60/141,581
; PRIOR FILING DATE: 1999-06-29
; NUMBER OF SEQ ID NOS: 5415
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 5006
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-606-776-5006

Query Match      98.9%; Score 1013; DB 26; Length 1998;
Best Local Similarity 99.9%; Pred. No. 1.5e-220;
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ACCGCTCGTCCGCGCGCTTGAAGCCCGGGAGCGC-GCGCAATTCGTGGCCCGCGG 59
Db 13 ACCGCTCGTCCGCGCGCTTGAAGCCCGGGAGCGCGGCGCAATTCGTGGCCCGCGG 72

QY 60 GGGGGGGGCTCCCGGCATCTTCGGCGGACCAAGGACTACCAGAAAGGGGCGGCTGG 119
Db 73 GGGGGGGGCTCCCGGCATCTTCGGCGGACCAAGGACTACCAGAAAGGGGCGGCTGG 132

QY 120 GATGGCGGCTCCGCGCGCGCGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 179
Db 133 GATGGCGGCTCCGCGCGCGCGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 192

QY 180 GGGCTACCCGCACTGGCGCGCGCGGATGTGAACTCCAGAGGGCGCTGTGAAGCCCTCC 239
Db 193 GGGCTACCCGCACTGGCGCGCGCGGATGTGAACTCCAGAGGGCGCTGTGAAGCCCTCC 252

QY 240 AGCAAAACAGTATCCTATCTCTTTTGGGCAACCATGAACTGCAATTCCTAGGTCCCAA 299
Db 253 AGCAAAACAGTATCCTATCTCTTTTGGGCAACCATGAACTGCAATTCCTAGGTCCCAA 312

QY 300 AGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACAAACGAAAGG 359
Db 313 AGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACAAACGAAAGG 372

QY 360 ATTTAACGAAGGATTTGGGAAATAGAAAATACCCAGAGTAAAGTTTACTGGCTACCA 419
Db 373 ATTTAACGAAGGATTTGGGAAATAGAAAATACCCAGAGTAAAGTTTACTGGCTACCA 432

QY 420 GCGAATTCAGCAACAGAGCTCTTCAGAACTCAGGAGAGAGTGAATATCTGCGATGTC 479
Db 433 GCGAATTCAGCAACAGAGCTCTTCAGAACTCAGGAGAGAGTGAATATCTGCGATGTC 492

QY 480 AAGCAGTGAAGAAAGGTTAGTAGTAGTGAAGAAAGTGAAGAAAGTGAAGAAAGTGA 539
Db 493 AAGCAGTGAAGAAAGGTTAGTAGTAGTGAAGAAAGTGAAGAAAGTGAAGAAAGTGA 552

QY 540 AAAAGCAGGCTCAAAACCGGAAAGTCAATATCTTCAAAGAAATCCTCTAAACAGTCCCG 599
Db 553 AAAAGCAGGCTCAAAACCGGAAAGTCAATATCTTCAAAGAAATCCTCTAAACAGTCCCG 612

QY 600 GAAATCTCAGAGATGAAGATGACAAAGACTGCAAGAGAGGAAACAAAGCAGCTC 659
Db 613 GAAATCTCAGAGATGAAGATGACAAAGACTGCAAGAGAGGAAACAAAGCAGCTC 672

QY 660 TGAGGCTGAGATGCGGGCAACGACACAAAGAAACAACTTCAGACTTTCAGAAAACAG 719
Db 673 TGAGGCTGAGATGCGGGCAACGACACAAAGAAACAACTTCAGACTTTCAGAAAACAG 732

QY 720 TGAAGGACCTTAACCTACCAATATGAATGCTGCATATTAAGAAACCAACAAAGAGTTAT 779
Db 733 TGAAGGACCTTAACCTACCAATATGAATGCTGCATATTAAGAAACCAACAAAGAGTTAT 792

QY 780 ATGTTTGGTGTCTAATATCTTGGATTTGATATGAACCAACACATAGTCTTGTGTCA 839
Db 793 ATGTTTGGTGTCTAATATCTTGGATTTGATATGAACCAACACATAGTCTTGTGTCA 852

QY 840 TTGACAGAACCCAGTTTGTATGTATATTTATTTATTTCTCTCTGTTGTGTTTCGGGG 899
Db 853 TTGACAGAACCCAGTTTGTATGTATATTTATTTATTTCTCTCTGTTGTGTTTCGGGG 912

QY 900 GAAAAGACATTTTAGCCTTTTAAAGGTTACTGATTTAATTTCAATTTATTTGTTGCA 959
Db 913 GAAAAGACATTTTAGCCTTTTAAAGGTTACTGATTTAATTTCAATTTATTTGTTGCA 972
```

```
QY 960 TGAAGTTGCCCTTAACCACTAAGATTATCAAGATTTTGGCAGACTTATACATGTCTA 1019
Db 973 TGAAGTTGCCCTTAACCACTAAGATTATCAAGATTTTGGCAGACTTATACATGTCTA 1032

QY 1020 GGATC 1024
Db 1033 GGATC 1037

RESULT 7
US-09-644-869-9226
; Sequence 9226, Application US/09644869
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Monahan, John
; APPLICANT: Richardson, Jennifer
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1182-001
; CURRENT APPLICATION NUMBER: US/09/644,869
; CURRENT FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: 60/151,062
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 9708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9226
; LENGTH: 2342
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-644-869-9226
```

```
Query Match      98.9%; Score 1013; DB 28; Length 2342;
Best Local Similarity 99.9%; Pred. No. 1.6e-220;
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
QY 1 ACCGCTCGTCCGCGCGCTTGAAGCCCGGGAGCGC-GCGCAATTCGTGGCCCGCGG 59
Db 13 ACCGCTCGTCCGCGCGCTTGAAGCCCGGGAGCGCGGCGCAATTCGTGGCCCGCGG 72

QY 60 GGGGGGGGCTCCCGGCATCTTCGGCGGACCAAGGACTACAGGAAGGGGAGCGGCTGG 119
Db 73 GGGGGGGGCTCCCGGCATCTTCGGCGGACCAAGGACTACAGGAAGGGGAGCGGCTGG 132

QY 120 GATGGCGGCTCCGCGCGCGCGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 179
Db 133 GATGGCGGCTCCGCGCGCGCGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 192

QY 180 GGGCTACCCGCACTGGCGCGCGCGGATGTGAACTCCAGAGGGCGCTGTGAAGCCCTCC 239
Db 193 GGGCTACCCGCACTGGCGCGCGCGGATGTGAACTCCAGAGGGCGCTGTGAAGCCCTCC 252

QY 240 AGCAAAACAGTATCCTATCTCTTTTGGGCAACCATGAAACTGCATTTCTAGGTCCCAA 299
Db 253 AGCAAAACAGTATCCTATCTCTTTTGGGCAACCATGAAACTGCATTTCTAGGTCCCAA 312

QY 300 AGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACAAACGAAAGG 359
Db 313 AGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACAAACGAAAGG 372

QY 360 ATTTAACGAAGGATTTGGGAAATAGAAAATACCCAGAGTAAAGTTTACTGGCTACCA 419
Db 373 ATTTAACGAAGGATTTGGGAAATAGAAAATACCCAGAGTAAAGTTTACTGGCTACCA 432

QY 420 GCGAATTCAGCAACAGAGCTCTTCAGAACTCAGGAGAGAGTGAATATCTGCGATGTC 479
Db 433 GCGAATTCAGCAACAGAGCTCTTCAGAACTCAGGAGAGAGTGAATATCTGCGATGTC 492

QY 480 AAGCAGTGAAGAAAGGTTAGTAGTAGTGAAGAAAGTGAAGAAAGTGAAGAAAGTGA 539
Db 493 AAGCAGTGAAGAAAGGTTAGTAGTAGTGAAGAAAGTGAAGAAAGTGAAGAAAGTGA 552
```

Qy 540 AAAAGCAGGCTCAAAAAGGAAAAAGTCAATATACCTTCAAGAAATCTCTTAAACAGTCCG 599
Db 553 AAAAGCAGGCTCAAAAAGGAAAAAGTCAATATACCTTCAAGAAATCTCTTAAACAGTCCG 612
Qy 600 GAAATCTCCAGGAGATGAAGATGCAAGAGTCAAGAGTCAAGAGAGGAAACAAACAGAGTCC 659
Db 613 GAAATCTCCAGGAGATGAAGATGCAAGAGTCAAGAGTCAAGAGAGGAAACAAACAGAGTCC 672
Qy 660 TGAGGCTGGAGATCGGGCAACGACACAAAGAAACCAACTTCAGACTTCAGAGAAACACAG 719
Db 673 TGAGGCTGGAGATCGGGCAACGACACAAAGAAACCAACTTCAGACTTCAGAGAAACACAG 732
Qy 720 TGAAGGACCTTAACCTAATGAATGCTCTCATATTAAGAGAAACCAAGAGAGTTAT 779
Db 733 TGAAGGACCTTAACCTAATGAATGCTCTCATATTAAGAGAAACCAAGAGAGTTAT 792
Qy 780 ATGTTGGTCTCTAATATCTTGGATTTGATATGAACCAACACATAGTCTTGTGCTCA 839
Db 793 ATGTTGGTCTCTAATATCTTGGATTTGATATGAACCAACACATAGTCTTGTGCTCA 852
Qy 840 TTGACAGAACCCAGTTTGTATGATATTAATTTTAAAGTTACTGATTTTAAATTTTGGTTCGA 959
Db 913 GAAAGACATTTAGCTCTTTTAAAGTTACTGATTTTAAATTTTAAATTTTGGTTCGA 972
Qy 960 TGAAGTTGCCCTTAACCACTAAGATTAATCAAGATTTTGGCAGACTTATACATGTCTA 1019
Db 973 TGAAGTTGCCCTTAACCACTAAGATTAATCAAGATTTTGGCAGACTTATACATGTCTA 1032
Qy 1020 GGATC 1024
Db 1033 GGATC 1037

RESULT 8

US-09-644-870-9669

; Sequence 9669, Application US/09644870

; GENERAL INFORMATION:

; APPLICANT: DiStefano, Peter

; APPLICANT: Holtzman, Douglas A.

; APPLICANT: Curtis, Rory A.J.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES

; TITLE OF INVENTION: THEREFOR

; FILE REFERENCE: 1600.1172-001

; CURRENT APPLICATION NUMBER: US/09/644,870

; CURRENT FILING DATE: 2000-08-28

; PRIOR APPLICATION NUMBER: 60/150,752

; PRIOR FILING DATE: 1999-08-26

; NUMBER OF SEQ ID NOS: 10434

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 9669

; LENGTH: 2342

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-644-870-9669

Query Match 98.9%; Score 1013; DB 28; Length 2342;
Best Local Similarity 99.9%; Pred. No. 1.6e-220;
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 1 ACCGCTGTCGCGCGGCTTGAAGCCCGCGGGAGCGG-GGCAATTCGTGCGCCGCGG 59
Db 13 ACCGCTGTCGCGCGGCTTGAAGCCCGCGGGAGCGCGCAATTCGTGCGCCGCGG 72
Qy 60 GGGGGCGGCTCCCGGATCTTCGCGGACCAAGGACTACCAAGAGGGAGCGGCTGG 119
Db 73 GGGGGCGGCTCCCGGATCTTCGCGGACCAAGGACTACCAAGAGGGAGCGGCTGG 132
Qy 120 GATGGCGGTCGCGCGGCGCGCGAGTACAAAGCGGGGACCTGGTCTTCGCCAAGATGA 179

Db 133 GATGGCGGTCGCGCGGCGCGAGTACAAAGCGGGGCGACCTGGTCTTCGCCAAGATGA 192
Qy 180 GGGCTACCCGCACTGGCGCGCGCGGATTAAGTAACTCCAGAGGGCGCTGTGAAGCTCC 239
Db 193 GGGCTACCCGCACTGGCGCGCGCGGATTAAGTAACTCCAGAGGGCGCTGTGAAGCTCC 252
Qy 240 AGCAAAACAAGTATCTATCTTTTGGCAGCCCATGAAACTGCACTTTCTAGGTCCAA 299
Db 253 AGCAAAACAAGTATCTATCTTTTGGCAGCCCATGCACTTTCTAGGTCCAA 312
Qy 300 AGACCTTTTCCATATAAGGAGTACAAGAACAAGTTTGGAAAGTCAAAACAAACGAAAGG 359
Db 313 AGACCTTTTCCATATAAGGAGTACAAGAACAAGTTTGGAAAGTCAAAACAAACGAAAGG 372
Qy 360 ATTTAAACGAAGGATTTGGGAAATAGAAAATACCCAGAGGAGTAAAGTTTACTGGCTACCA 419
Db 373 ATTTAAACGAAGGATTTGGGAAATAGAAAATACCCAGGAGTAAAGTTTACTGGCTACCA 432
Qy 420 GGCAATTTCAAGCAACAGAGCTCTTTCAGAAACTGAGGGAGAGGTGGAAATATCTGCAGATGC 479
Db 433 GGCAATTTCAAGCAACAGAGCTCTTTCAGAAACTGAGGGAGAGGTGGAAATATCTGCAGATGC 492
Qy 480 AAGCAGTGAAGAAAGAGTGTATAGATAGAGAGATGGAAGAGGCAAAAGAAAGATGA 539
Db 493 AAGCAGTGAAGAAAGAGTGTATAGATAGAGAGATGGAAGAGGCAAAAGAAAGATGA 552
Qy 540 AAAAGCAGGCTCAAAAAGGAAAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCG 599
Db 553 AAAAGCAGGCTCAAAAAGGAAAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCG 612
Qy 600 GAAATCTCCAGGAGATGAAGATGCAAGACTGCAAGAGAGAGGAAACAAACAGCAGTCC 659
Db 613 GAAATCTCCAGGAGATGAAGATGCAAGACTGCAAGAGAGAGGAAACAAACAGCAGTCC 672
Qy 660 TGAGGCTGGAGATCGGGCAACGACACAAAGAAACCAACTTCAGACTTCGAGAAACACAG 719
Db 673 TGAGGCTGGAGATCGGGCAACGACACAAAGAAACCAACTTCAGACTTCGAGAAACACAG 732
Qy 720 TGAAGGACCTTAACCTAATGAATGCTGCAATTAAGAGAAACCAAGAGAGTTAT 779
Db 733 TGAAGGACCTTAACCTAATGAATGCTGCAATTAAGAGAAACCAAGAGAGTTAT 792
Qy 780 ATGTTGGTCTCTAATATCTTGGATTTGATATGAACCAACACATAGTCTTGTGCTCA 839
Db 793 ATGTTGGTCTCTAATATCTTGGATTTGATATGAACCAACACATAGTCTTGTGCTCA 852
Qy 840 TTGACAGAACCCAGTTTGTATGATTAATTTTAAATTTCTCTCTGTGTGTTTCGGGG 899
Db 853 TTGACAGAACCCAGTTTGTATGATTAATTTTAAATTTCTCTCTGTGTGTTTCGGGG 912
Qy 900 GAAAGACATTTTAGCTCTTTTAAAGTTACTGATTTTAAATTTTAAATTTTGGTTCGA 959
Db 913 GAAAGACATTTTAGCTCTTTTAAAGTTACTGATTTTAAATTTTAAATTTTGGTTCGA 972
Qy 960 TGAAGTTGCCCTTAACCACTAAGATTAATCAAGATTTTGGCAGACTTATACATGTCTA 1019
Db 973 TGAAGTTGCCCTTAACCACTAAGATTAATCAAGATTTTGGCAGACTTATACATGTCTA 1032
Qy 1020 GGATC 1024
Db 1033 GGATC 1037

RESULT 9

US-09-649-162-9840

; Sequence 9840, Application US/09649162

; GENERAL INFORMATION:

; APPLICANT: Shyjan, Andrew W.

; APPLICANT: Richardson, Jennifer

; APPLICANT: Holtzman, Douglas A.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES

; TITLE OF INVENTION: THEREFOR

```
; FILE REFERENCE: 1600.1181-001
; CURRENT APPLICATION NUMBER: US/09/649,162
; CURRENT FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: 60/151,057
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 9990
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9840
; LENGTH: 2342
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-649-162-9840

Query Match      98.9%; Score 1013; DB 28; Length 2342;
Best Local Similarity 99.9%; Pred. No. 1.6e-220;
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ACCGCTGTCGCCCGCCGCTTACGCGCGGAGCGC-GCGCAATTCGTCGCGCGCGG 59
Db 13 ACCGCTGTCGCCCGCCGCTTACGCGCGGAGCGCGCAATTCGTCGCGCGCGG 72
QY 60 GGGGGCGGCTCCCGCGCTTCGCGCGGACCAAGACTACCAGGAGGAGCGGCTGG 119
Db 73 GGGGGCGGCTCCCGCGCTTCGCGCGGACCAAGACTACCAGGAGGAGCGGCTGG 132
QY 120 GATGGCGGCTCCCGCGCGCGGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 179
Db 133 GATGGCGGCTCCCGCGCGCGGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 192
QY 180 GGGCTACCCGCACTGGCGCGCGCGGATGATGAACTCCAGAGGCGGCTGTGAAGCTCC 239
Db 193 GGGCTACCCGCACTGGCGCGCGCGGATGATGAACTCCAGAGGCGGCTGTGAAGCTCC 252
QY 240 AGCAATTCAGACAGCTCTTCAGAACTGAGGAGAGGTGGAAATCTCTAAACAGTCCG 479
Db 433 GGCAATTCAGACAGCTCTTCAGAACTGAGGAGAGGTGGAAATCTCTAAACAGTCCG 492
QY 480 AAGCAGTGGAGGAGAGGTGATAGAGTACAAAGAGATGAAAGCGCAAAAGAAAGATGA 539
Db 493 AAGCAGTGGAGGAGAGGTGATAGAGTACAAAGAGATGAAAGCGCAAAAGAAAGATGA 552
QY 540 AAAAGCAGGCTCAAAACGGAAGAGTATATATCTTCAAGAGAAATCTCTAAACAGTCCG 599
Db 553 AAAAGCAGGCTCAAAACGGAAGAGTATATATCTTCAAGAGAAATCTCTAAACAGTCCG 612
QY 600 GAAATCTCAGGAGATGATGATGACAAAGACTGCAAGAGAGGAGGAAACAAAGCAGCTC 659
Db 613 GAAATCTCAGGAGATGATGATGACAAAGACTGCAAGAGAGGAGGAAACAAAGCAGCTC 672
QY 660 TGAGGTGGAGATGGCGGCAACGACACAAAGAACCAACTTCAGACTTCGAGAAACACAG 719
Db 673 TGAGGTGGAGATGGCGGCAACGACACAAAGAACCAACTTCAGACTTCGAGAAACACAG 732
QY 720 TGAAGGGAACCTAATACCAATGATGCTGCATATTAAGAGAGAAACCAAGAGGTTAT 779
Db 733 TGAAGGGAACCTAATACCAATGATGCTGCATATTAAGAGAGAAACCAAGAGGTTAT 792
QY 780 ATGTTTGGTGTCTAATATCTTGGATTTGATATGAAACCAACATAGTCCCTTGTGTCA 839
Db 793 ATGTTTGGTGTCTAATATCTTGGATTTGATATGAAACCAACATAGTCCCTTGTGTCA 852
```

```
QY 840 TTGACAGAACCCAGTTGTATGTATCATATTATTCATATTCCTCTCTGTCTGTTCGGGG 899
Db 853 TTGACAGAACCCAGTTGTATGTATCATATTATTCATATTCCTCTCTGTCTGTTCGGGG 912
QY 900 GAAAAGACATTTTAGCCTTTTAAAGTTACTGATTTAAATTCATTTTGGTTGCA 959
Db 913 GAAAAGACATTTTAGCCTTTTAAAGTTACTGATTTAAATTCATTTTGGTTGCA 972
QY 960 TGAAGTTGCCCTTAACCACTAAGGATTATCAAGATTTTGGCGAGACTTATACATGCTA 1019
Db 973 TGAAGTTGCCCTTAACCACTAAGGATTATCAAGATTTTGGCGAGACTTATACATGCTA 1032
QY 1020 GGATC 1024
Db 1033 GGATC 1037

RESULT 10
US-09-652-109-9938
; Sequence 9938, Application US/09652109
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1180-001
; CURRENT APPLICATION NUMBER: US/09/652,109
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,128
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 10105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9938
; LENGTH: 2342
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-109-9938

Query Match      98.9%; Score 1013; DB 28; Length 2342;
Best Local Similarity 99.9%; Pred. No. 1.6e-220;
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ACCGCTGTCGCCCGCCGCTTACGCGCGGAGCGC-GCGCAATTCGTCGCGCGCGG 59
Db 13 ACCGCTGTCGCCCGCCGCTTACGCGCGGAGCGCGCAATTCGTCGCGCGCGG 72
QY 60 GGGGGCGGCTCCCGCGCTTCGCGCGGACCAAGACTACCAGGAGGAGCGGCTGG 119
Db 73 GGGGGCGGCTCCCGCGCTTCGCGCGGACCAAGACTACCAGGAGGAGCGGCTGG 132
QY 120 GATGGCGGCTCCCGCGCGCGGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 179
Db 133 GATGGCGGCTCCCGCGCGCGGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 192
QY 180 GGGCTACCCGCACTGGCGCGCGCGGATGATGAACTCCAGAGGCGGCTGTGAAGCTCC 239
Db 193 GGGCTACCCGCACTGGCGCGCGCGGATGATGAACTCCAGAGGCGGCTGTGAAGCTCC 252
QY 240 AGCAATTCAGACAGCTCTTCAGAACTGAGGAGAGGTGGAAATCTCTAAACAGTCCG 299
Db 253 AGCAATTCAGACAGCTCTTCAGAACTGAGGAGAGGTGGAAATCTCTAAACAGTCCG 312
QY 300 AGACCTTTTCCATATAAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACCGGAAAGG 359
Db 313 AGACCTTTTCCATATAAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACCGGAAAGG 372
QY 360 ATTTAACGAGAGGATTTGGGAAATAGAAAATAACCCAGAGTAAAGTTTCTGGCTACCA 419
Db 373 ATTTAACGAGAGGATTTGGGAAATAGAAAATAACCCAGAGTAAAGTTTCTGGCTACCA 432
QY 420 GGCATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTGGAAATCTCTAAACAGTCC 479
Db 433 GGCATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTGGAAATCTCTAAACAGTCC 492
```


; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.2013-001
; CURRENT APPLICATION NUMBER: US/09/698,013
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 60/162,360
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 7935
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6892
; LENGTH: 2342
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-698-013-6892

Query Match 98.9%; Score 1013; DB 30; Length 2342;
Best Local Similarity 99.9%; Pred. No. 1.6e-220;
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ACCGCTCGTCCGCCCGCTTGAAGCCCGCGGGAGCGC-GCGCAATTCTGTCGCCCGCGG 59
Db 13 ACCGCTCGTCCGCCCGCTTGAAGCCCGCGGGAGCGCGCAATTCTGTCGCCCGCGG 72

QY 60 GGGGGCGGCTCCCGGCACTCTTCCGGCGACCAAGGACTACAGGAAGGGAGCGGCTGG 119
Db 73 GGGGGCGGCTCCCGGCACTCTTCCGGCGACCAAGGACTACAGGAAGGGAGCGGCTGG 132

QY 120 GATGGCGGCTCCCGGCCCGCGAGTACAAAGCGGCGACCTGCTTCCGCCAAGATGA 179
Db 133 GATGGCGGCTCCCGGCCCGCGAGTACAAAGCGGCGACCTGCTTCCGCCAAGATGA 192

QY 180 GGGCTACCCGCACTGGCCCGCGCGGCTTGAAGCCCGCGAGGAGCGGCTGG 239
Db 193 GGGCTACCCGCACTGGCCCGCGCGGCTTGAAGCCCGCGAGGAGCGGCTGG 252

QY 240 AGCAACCAAGTATCCTATCTTCTTTTGGCACCCATGAACTGATTTCTAGGTCCTCA 299
Db 253 AGCAACCAAGTATCCTATCTTCTTTTGGCACCCATGAACTGATTTCTAGGTCCTCA 312

QY 300 AGACCTTTTCCATATAAGGAGTACAAAGACCAAGTTCGAAAGTCAAAACGAAAGG 359
Db 313 AGACCTTTTCCATATAAGGAGTACAAAGACCAAGTTCGAAAGTCAAAACGAAAGG 372

QY 360 ATTTAACGAAGGATTGTGGAAATAGAAATACCCAGGATTAAGTTTACTGGCTACCA 419
Db 373 ATTTAACGAAGGATTGTGGAAATAGAAATACCCAGGATTAAGTTTACTGGCTACCA 432

QY 420 GGCATTTCCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTTGAAATACTGCAAGTGC 479
Db 433 GGCATTTCCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTTGAAATACTGCAAGTGC 492

QY 480 AAGCAGTGAAGAGGATGATAGTAGTAGAAGATGAAAGGAGGAAAGCAAGATGA 539
Db 493 AAGCAGTGAAGAGGATGATAGTAGTAGAAGATGAAAGGAGGAAAGCAAGATGA 552

QY 540 AAAAGCAGGCTCAAAACGAAAGTCTATATCTTCAAGAACTCTTCAAGAGTCTTCA 599
Db 553 AAAAGCAGGCTCAAAACGAAAGTCTATATCTTCAAGAACTCTTCAAGAGTCTTCA 612

QY 600 GAAATCTCCAGGAGTGAAGTACAAAGACTGCAAGAGGAGGAGGAAACAAAGCAGCTC 659
Db 613 GAAATCTCCAGGAGTGAAGTACAAAGACTGCAAGAGGAGGAGGAAACAAAGCAGCTC 672

QY 660 TGAGGTTGAGATGCGGGCAACGACACAAAGAAACCACTTCCAGACTTGCAGAAACCG 719
Db 673 TGAGGTTGAGATGCGGGCAACGACACAAAGAAACCACTTCCAGACTTGCAGAAACCG 732

QY 720 TCAAGGACCTAACTACCAATATGATCTGCTATTAAGGAAACCAACCAAGAGTTAT 779
Db 733 TCAAGGACCTAACTACCAATATGATCTGCTATTAAGGAAACCAACCAAGAGTTAT 792

QY 780 ATGTTTGGTGTCTAATATCTTGGATTGATGAACCAACACATAGTCTTGTGTCA 839

Db 793 ATGTTTGGTGTCTAATATCTTGGATTGATGAACCAACACATAGTCTTGTGTCA 852
QY 840 TTGACAGAAACCCAGTTTGTATGATACATTATTATTTATTTCTCTGTTGTGTTTGGGG 899
Db 853 TTGACAGAAACCCAGTTTGTATGATACATTATTATTTCTCTCTGTTGTGTTTGGGG 912
QY 900 GAAAGACATTTTAGCCCTTTTAAAAAGTTACTGATTTTAAATTTATTTATTTGGTTGCA 959
Db 913 GAAAGACATTTTAGCCCTTTTAAAAAGTTACTGATTTTAAATTTATTTATTTGGTTGCA 972
QY 960 TGAAGTGGCTTTAACCACTAAGGATTTAAGATTTTGGCAGACTTATACATGCTTA 1019
Db 973 TGAAGTGGCTTTAACCACTAAGGATTTAAGATTTTGGCAGACTTATACATGCTTA 1032
QY 1020 GGATC 1024
Db 1033 GGATC 1037

RESULT 13
US-09-726-175-3077
; Sequence 3077, Application US/09726175
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Fraser, Christopher C.
; APPLICANT: Donovan, Michael J.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2054-001
; CURRENT APPLICATION NUMBER: US/09/726,175
; CURRENT FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/167,859
; PRIOR FILING DATE: 1999-11-29
; NUMBER OF SEQ ID NOS: 3770
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3077
; LENGTH: 2342
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-726-175-3077

Query Match 98.9%; Score 1013; DB 32; Length 2342;
Best Local Similarity 99.9%; Pred. No. 1.6e-220;
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ACCGCTCGTCCGCCCGCTTGAAGCCCGCGGGAGCGC-GCGCAATTCTGTCGCCCGCGG 59
Db 13 ACCGCTCGTCCGCCCGCTTGAAGCCCGCGGGAGCGCGCAATTCTGTCGCCCGCGG 72

QY 60 GGGGGCGGCTCCCGGCACTCTTCCGGCGACCAAGGACTACAGGAAGGGAGCGGCTGG 119
Db 73 GGGGGCGGCTCCCGGCACTCTTCCGGCGACCAAGGACTACAGGAAGGGAGCGGCTGG 132

QY 120 GATGGCGGCTCCCGGCCCGCGAGTACAAAGCGGCGACCTGCTTCCGCCAAGATGA 179
Db 133 GATGGCGGCTCCCGGCCCGCGAGTACAAAGCGGCGACCTGCTTCCGCCAAGATGA 192

QY 180 GGGCTACCCGCACTGGCCCGCGCGGCTTGAAGCCCGCGAGGAGCGGCTGG 239
Db 193 GGGCTACCCGCACTGGCCCGCGCGGCTTGAAGCCCGCGAGGAGCGGCTGG 252

QY 240 AGCAACCAAGTATCCTATCTTCTTTTGGCACCCATGAACTGATTTCTAGGTCCTCA 299
Db 253 AGCAACCAAGTATCCTATCTTCTTTTGGCACCCATGAACTGATTTCTAGGTCCTCA 312

QY 300 AGACCTTTTCCATATAAGGAGTACAAAGACCAAGTTCGAAAGTCAAAACGAAAGG 359
Db 313 AGACCTTTTCCATATAAGGAGTACAAAGACCAAGTTCGAAAGTCAAAACGAAAGG 372

QY 360 ATTTAACGAAGGATTGTGGAAATAGAAATACCCAGGATTAAGTTTACTGGCTACCA 419
Db 373 ATTTAACGAAGGATTGTGGAAATAGAAATACCCAGGATTAAGTTTACTGGCTACCA 432

QY 420 GGCATTTTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAAAGTGGAAATACTGCAGATGC 479
DB 433 GGCATTTTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAAAGTGGAAATACTGCAGATGC 492
QY 480 AACGAGTGAGGAAGAGTGTATAGATGATAGAGAGATGGAAGAGGCAAAAGAAAGATGA 539
DB 493 AACGAGTGAGGAAGAGTGTATAGATGATAGAGAGATGGAAGAGGCAAAAGAAAGATGA 552
QY 540 AAAAGCAGGCTCAAAACCGAAAAGTCTATATCTTCAAGAAATCTCTTAAACAGATCCCG 599
DB 553 AAAAGCAGGCTCAAAACCGAAAAGTCTATATCTTCAAGAAATCTCTTAAACAGATCCCG 612
QY 600 GAAATCTCCAGGAGATGAAGATGACAAAGACTGCAAAAGAGAGGAAACAAAAGCAGCTC 659
DB 613 GAAATCTCCAGGAGATGAAGATGACAAAGACTGCAAAAGAGAGGAAACAAAAGCAGCTC 672
QY 660 TGAGGTTGGAGATGCGGGCAACGACACAGAAACAACTTCAGACTTGCAGAAAACCCAG 719
DB 673 TGAGGTTGGAGATGCGGGCAACGACACAGAAACAACTTCAGACTTGCAGAAAACCCAG 732
QY 720 TGAAGGACCTAACTACCATATGAATGCTGCTGATATTAAGAGAAACCCACAGAAAGTGTAT 779
DB 733 TGAAGGACCTAACTACCATATGAATGCTGCTGATATTAAGAGAAACCCACAGAAAGTGTAT 792
QY 780 ATGTTTGGTGTCTAATATTTCTGGATTTGATATGAACCAACACATAGTCTCTTGTGTCA 839
DB 793 ATGTTTGGTGTCTAATATTTCTGGATTTGATATGAACCAACACATAGTCTCTTGTGTCA 852
QY 840 TTGACAGAACCCAGATTTGTATGATATTAATTTCTCTCTCTGTTGTTTGGGGG 899
DB 853 TTGACAGAACCCAGATTTGTATGATATTAATTTCTCTCTCTGTTGTTTGGGGG 912
QY 900 GAAAGACATTTTAGCCCTTTTAAAGTTTACTGATTTAATTTCAATGTTTGGTTGCA 959
DB 913 GAAAGACATTTTAGCCCTTTTAAAGTTTACTGATTTAATTTCAATGTTTGGTTGCA 972
QY 960 TGAAGTTGCCCTTAACCACTAAGGATTTATCAAGATTTTTCGCGAGACTTATACATGTCTA 1019
DB 973 TGAAGTTGCCCTTAACCACTAAGGATTTATCAAGATTTTTCGCGAGACTTATACATGTCTA 1032
QY 1020 GGATC 1024
DB 1033 GGATC 1037

RESULT 14

US-09-726-806-5292
; Sequence 5292, Application US/09726806
; GENERAL INFORMATION:
; APPLICANT: Galvin, Katherine
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2028-001
; CURRENT APPLICATION NUMBER: US/09/726.806
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/168.135
; PRIOR FILING DATE: 1999-11-30
; NUMBER OF SEQ ID NOS: 6283
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5292
; LENGTH: 2342
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-726-806-5292

Query Match 98.9%; Score 1013; DB 32; Length 2342;
Best Local Similarity 99.9%; Pred. No. 1.6e-220;
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 ACCGCTCGTCCGCGCTTTAGGCGCGCGGGAGCGC-GCGCAATTCGTCGCGCGCGG 59
DB 13 ACCGCTCGTCCGCGCTTTAGGCGCGCGGGAGCGCGCAATTCGTCGCGCGCGG 72

QY 60 GGGGGCGGCTCCCGGCTCTTCGCGCATCTTCGCGCGACCAAGGACTACCAAGAGGGAGCGGCTGG 119
DB 73 GGGGGCGGCTCCCGGCTCTTCGCGCATCTTCGCGCGACCAAGGACTACCAAGAGGGAGCGGCTGG 132
QY 120 GATGGCGGCTCCCGGCGCGCGGAGTACAAAGCGGCGGAGCTTGGTCTTCCGCAAGATGAA 179
DB 133 GATGGCGGCTCCCGGCGCGCGGAGTACAAAGCGGCGGAGCTTGGTCTTCCGCAAGATGAA 192
QY 180 GGGCTACCCGCGACTGCGCGCGCGGATTCAGTAACCTCCAGAGGGCGCTGTGAAGCTCC 239
DB 193 GGGCTACCCGCGACTGCGCGCGCGGATTCAGTAACCTCCAGAGGGCGCTGTGAAGCTCC 252
QY 240 AGCAAAACAAGTATCTTCTTTTGGCACCCTATGAACTGCAATTTCTAGGCTCCCA 299
DB 253 AGCAAAACAAGTATCTTCTTTTGGCACCCTATGAACTGCAATTTCTAGGCTCCCA 312
QY 300 AGACCTTTTTCATATAAGAGGTACAAAGACAAAGTTTGGAAAGTCAAAACGGAAGG 359
DB 313 AGACCTTTTTCATATAAGAGGTACAAAGACAAAGTTTGGAAAGTCAAAACGGAAGG 372
QY 360 ATTTRAAGAGGATTTGGGGAATAGAAATACCCAGGAGTAAAGTTTACTGGCTACCA 419
DB 373 ATTTRAAGAGGATTTGGGGAATAGAAATACCCAGGAGTAAAGTTTACTGGCTACCA 432
QY 420 GGCATTTTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAAAGTGGAAATACTGCAGATGC 479
DB 433 GGCATTTTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAAAGTGGAAATACTGCAGATGC 492
QY 480 AAGCAGTGAGGAAGAGTGTATAGATGATAGAGAGATGGAAGAGGCAAAAGAAAGATGA 539
DB 493 AAGCAGTGAGGAAGAGTGTATAGATGATAGAGAGATGGAAGAGGCAAAAGAAAGATGA 552
QY 540 AAAAGCAGGCTCAAAACCGAAAAGTCTATATCTTCAAGAAATCTCTTAAACAGTCCCG 599
DB 553 AAAAGCAGGCTCAAAACCGAAAAGTCTATATCTTCAAGAAATCTCTTAAACAGTCCCG 612
QY 600 GAAATCTCCAGGAGATGAAGATGACAAAGACTGCAAAAGAGAGGAAACAAAAGCAGCTC 659
DB 613 GAAATCTCCAGGAGATGAAGATGACAAAGACTGCAAAAGAGAGGAAACAAAAGCAGCTC 672
QY 660 TGAGGTTGGAGATGCGGGCAACGACACAAAGAAACAACTTCAGACTTGCAGAAAACCCAG 719
DB 673 TGAGGTTGGAGATGCGGGCAACGACACAAAGAAACAACTTCAGACTTGCAGAAAACCCAG 732
QY 720 TGAAGGACCTAACTACCATATGAATGCTGCTGATATTAAGAGAAACCCACAGAAAGTGTAT 779
DB 733 TGAAGGACCTAACTACCATATGAATGCTGCTGATATTAAGAGAAACCCACAGAAAGTGTAT 792
QY 780 ATGTTTGGTGTCTAATATTTCTGGATTTGATATGAACCAACACATAGTCTCTTGTGTCA 839
DB 793 ATGTTTGGTGTCTAATATTTCTGGATTTGATATGAACCAACACATAGTCTCTTGTGTCA 852
QY 840 TTGACAGAACCCAGATTTGTATGATATTAATTTCTCTCTCTGTTGTTTGGGGG 899
DB 853 TTGACAGAACCCAGATTTGTATGATATTAATTTCTCTCTCTGTTGTTTGGGGG 912
QY 900 GAAAGACATTTTAGCCCTTTTAAAGTTTACTGATTTAATTTCAATGTTTGGTTGCA 959
DB 913 GAAAGACATTTTAGCCCTTTTAAAGTTTACTGATTTAATTTCAATGTTTGGTTGCA 972
QY 960 TGAAGTTGCCCTTAACCACTAAGGATTTATCAAGATTTTTCGCGAGACTTATACATGTCTA 1019
DB 973 TGAAGTTGCCCTTAACCACTAAGGATTTATCAAGATTTTTCGCGAGACTTATACATGTCTA 1032
QY 1020 GGATC 1024
DB 1033 GGATC 1037

RESULT 15
US-09-770-173-2543
; Sequence 2543, Application US/09770173

; GENERAL INFORMATION:
; APPLICANT: White, David
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR

; FILE REFERENCE: 1600.2059-001
; CURRENT APPLICATION NUMBER: US/09/770,173
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,876
; PRIOR FILING DATE: 2000-01-28
; NUMBER OF SEQ ID NOS: 3167
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2543

; LENGTH: 1909
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-770-173-2543

Query Match 98.8%; Score 1012; DB 33; Length 1909;
Best Local Similarity 99.9%; Pred. No. 2.5e-220;
Matches 1023; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
QY 2 CCGCTCGTCGCGCGCGCTTGAGCGCGCGGAGCGC-GCGCAATTCGTCGCGCGCGCGG 60
Db |||||
QY 61 GGGGCGGCTCCGCGCATCTTCGCGGACCAAGGACTACCAAGGAGGGGCGGCTGG 120
Db |||||
QY 94 GGGGCGGCTCCGCGCATCTTCGCGGACCAAGGACTACCAAGGAGGGGCGGCTGG 153
Db |||||
QY 121 ATGGGCGGCTCCGCGCGCGCGAGTACAAAGCGGCGACCTCGTCTTCGCCAAGATGAAG 180
Db |||||
QY 154 ATGGCGGCTCCGCGCGCGCGAGTACAAAGCGGCGGCGCTCGTCTTCGCCAAGATGAAG 213
Db |||||
QY 181 GGTCTACCGCATCGCGCGCGCGGATGTATGAATCTCCAGAGGCGCTGTGAAGCTCCA 240
Db |||||
QY 214 GGTCTACCGCATCGCGCGCGCGGATGTATGAATCTCCAGAGGCGCTGTGAAGCTCCA 273
Db |||||
QY 241 GCAACACAGTATCCTATCTCTTTTGGCACCCATGAACTGCACTTTCTAGTCCCAA 300
Db |||||
QY 274 GCAACACAGTATCCTATCTCTTTTGGCACCCATGAACTGCACTTTCTAGTCCCAA 333
Db |||||
QY 301 GACCTTTTCCATATAAGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAGGA 360
Db |||||
QY 334 GACCTTTTCCATATAAGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAGGA 393
Db |||||
QY 361 TTTAAGAAAGGATTTGGGAAATAGAAATAACCCAGGAGTAAAGTTTACTCGCTACCG 420
Db |||||
QY 394 TTTAAGAAAGGATTTGGGAAATAGAAATAACCCAGGAGTAAAGTTTACTCGCTACCG 453
Db |||||
QY 421 GCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGGAGAGTGGAAATACTGCAGATGCA 480
Db |||||
QY 454 GCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGGAGAGTGGAAATACTGCAGATGCA 513
Db |||||
QY 481 AGCAGTGAGGAAAGAGTGATAGAGTAGAAGATGGAAGGCAAGAAAGAAAGATGAA 540
Db |||||
QY 514 AGCAGTGAGGAAAGAGTGATAGAGTAGAAGATGGAAGGCAAGAAAGAAAGATGAA 573
Db |||||
QY 541 AAAGCAGGCTCAAAACGGAAGAGTATATCTTCAAGAAATCTCTTAAACAGTCCCGG 600
Db |||||
QY 574 AAAGCAGGCTCAAAACGGAAGAGTATATCTTCAAGAAATCTCTTAAACAGTCCCGG 633
Db |||||
QY 601 AAATCTCCAGGAGTGAAGTACAAAGACTGCAAGAGAGGAAACAAAGAGCAGCTCT 660
Db |||||
QY 634 AAATCTCCAGGAGTGAAGTACAAAGACTGCAAGAGAGGAAACAAAGAGCAGCTCT 693
Db |||||
QY 661 GAGGCTGAGATGCGGGCAACCAACAGAAACAACTTCAGACTTCAGAAACCAAGT 720
Db |||||
QY 694 GAGGCTGAGATGCGGGCAACCAACAGAAACAACTTCAGACTTCAGAAACCAAGT 753
Db |||||
QY 721 GAAGGACCTTAACCTACCAATGAATGCTGCATATTTAAGAGAAACCAAGAGGTTATA 780
Db |||||
QY 754 GAAGGACCTTAACCTACCAATGAATGCTGCATATTTAAGAGAAACCAAGAGGTTATA 813
Db |||||
```

```
QY 781 TGTTCGTTGCTAATATTTCTTGGATTTGATATGAACCAACACATAGTCTTGTGTGTCAT 840
Db |||||
QY 814 TGTTCGTTGCTAATATTTCTTGGATTTGATATGAACCAACACATAGTCTTGTGTGTCAT 873
Db |||||
QY 841 TGACAGAAACCCAGTTTGTATGTACATTTATTCATATTCCTCTCTGTTGTGTTCCGGGGG 900
Db |||||
QY 874 TGACAGAAACCCAGTTTGTATGTACATTTATTCATATTCCTCTCTGTTGTGTTCCGGGGG 933
Db |||||
QY 901 AAAAGACATTTAGCCTTTTAAAGTTACTGATTTAAATTTCAATGTTTGTGTTGTCAT 960
Db |||||
QY 934 AAAAGACATTTAGCCTTTTAAAGTTACTGATTTAAATTTCAATGTTTGTGTTGTCAT 993
Db |||||
QY 961 GAAAGTGGCCCTTAACCACTAAGATTATCAAGATTTTGGCGCAGACTTATACATGCTCTAG 1020
Db |||||
QY 994 GAAAGTGGCCCTTAACCACTAAGATTATCAAGATTTTGGCGCAGACTTATACATGCTCTAG 1053
Db |||||
QY 1021 GATC 1024
Db 1054 GATC 1057
```

Search completed: January 18, 2004, 12:01:11
Job time : 3670 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2004, 08:34:45 ; Search time 2527 Seconds
(without alignments)
9848.747 Million cell updates/sec

Title: US-09-787-328B-3
Perfect score: 1024
Sequence: 1 accgcgtccgcgcgcgtt.....acttatacatgtctaggatc 1024

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estm:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_esti:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1007.2	98.4	2327	11 BC040554	BC040554 Homo sapi
2	977.8	95.5	1201	9 AL520288	AL520288 AL520288
3	907	88.6	1201	9 AL520469	AL520469 AL520469
4	764	74.6	808	10 BG201738	BG201738 RST21080

5	763.2	74.5	1112	12	BM457409	BM457409 AGENCOURT
6	758	74.0	1868	11	AK044858	Mus muscu
7	757.8	74.0	889	14	CD251499	AGENCOURT
8	754.4	73.7	2722	11	AK019487	Mus muscu
9	746	72.9	1526	11	AK014849	Mus muscu
10	746	72.9	3363	11	AK045542	Mus muscu
11	739	72.2	819	10	BG204817	RST24232
12	690.4	67.4	692	12	BM722135	UI-E-E00-
13	672.2	65.6	781	10	BE886964	601506447
14	659.6	64.4	813	10	BG212737	RST32332
15	653	63.8	796	14	CA315695	UI-M-FW0-
16	639.6	62.5	772	14	CD000122	AGENCOURT
17	631	61.6	643	14	CB215502	NISC np06
18	630.8	61.6	919	10	BF346836	602021638
19	623.6	60.9	737	9	AU126115	AU126115
20	613.2	59.9	914	10	BF306828	601891619
21	605.2	59.1	681	10	BG719650	602689854
22	597.6	58.4	664	12	B1400803	MI-P-AY1-
23	593.6	58.0	2788	11	AK051401	Mus muscu
24	576.8	56.3	671	14	BY721555	BY721555
25	572.8	55.9	707	12	B1202384	UI-M-CG09
26	568	55.5	863	13	BUI79069	AGENCOURT
27	562.6	54.9	876	13	BUI64599	AGENCOURT
28	562.4	54.9	580	9	AI743609	wg51f12.x
29	552.8	54.0	557	14	CA941241	1f32602.x
30	551.8	53.9	651	14	CD354909	UI-M-GM0-
31	549.4	53.7	663	10	BE111763	UI-R-BJ1-
32	543	53.0	559	10	BE703369	MR2-NN111
33	539	52.6	566	10	BE703403	MR2-NN111
C	536.2	52.4	654	10	BE703421	MR2-NN111
35	527.2	51.5	532	10	BF445986	7p17d09.x
36	525.6	51.3	557	10	BE348635	htc73c04.x
C	525.4	51.3	547	12	BM931689	UI-R-EJ1-
38	522.8	51.1	973	13	BQ685483	AGENCOURT
C	522.4	51.0	545	12	BM931237	UI-E-EJ1-
C	522.4	51.0	545	13	BQ186294	UI-E-EJ1-
41	522.4	51.0	560	10	BE550171	7b49g08.x
C	519.6	50.7	538	12	BM725299	UI-R-EJ0-
C	514.2	50.2	531	9	AW161910	au71e02.x
C	513.6	50.2	608	10	BE703418	MR2-NN111
C	513	50.1	602	10	BE703252	MR2-NN111

ALIGNMENTS

RESULT 1
BC040554

LOCUS
DEFINITION

BC040554
Homo sapiens, similar to likely ortholog of mouse hepatoma-derived growth factor, related protein 3, clone IMAGE:5303793, mRNA.

ACCESSION
BC040554

VERSION
BC040554.1

KEYWORDS
HTC

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 2327)

AUTHORS
Straussberg, R.

TITLE
Direct Submission

JOURNAL
Submitted (29-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
Contact: MGC help desk

COMMENT
Email: cgaops@mail.nih.gov

BC040554 2327 bp mRNA linear HTC 04-MAR-2003
Homo sapiens, similar to likely ortholog of mouse hepatoma-derived growth factor, related protein 3, clone IMAGE:5303793, mRNA.
BC040554 GI:26996855
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2327)
Direct Submission
Submitted (29-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgaops@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome

QY	CGCGTCGTCGCGCCGGCTTGAGGCCCGGGGAGCGC-GGCAATTTCGTTCGGCCCGCGGG	60
488	CGCGTCGTCGCGCCGGCTTGAGGCCCGGGGAGCGCGGCGCAATTCGTTCGGCCCGCGGG	547
QY	GGGGCGCGCTCCCGGCATCTTCGCGGCGACCAAGGACTACAGGAAGGGGAGCGGCTGGG	120
Db	GGGGCGCGCTCCCGGCATCTTCAGGCGACCAAGGACTACAGGAAGGGGAGCGGCTGGG	607
QY	ATGGCGCGTCCGGCGCCCGCGAGTACAAAGCGGGCGACCTGGTCTTCGCGCAAGATGAAG	180
Db	ATGGCGCGTCCGGCGCCCGCGAGTACAAAGCGGGCGACCTGGTCTTCGCGCAAGATGAAG	667
QY	GGCTACCGCGACTGGCGGCCCGGATGTGTAACCTCCAGAGGGCGCTGTGAAGCCTCCA	240
Db	GGCTACCGCGACTGGCGGCCCGGATGTGTAACCTCCAGAGGGCGCTGTGAAGCCTCCA	727
QY	GCAACCAAGATATCTATCTCTCTTTTGGCACCCATGAAACTGCATTTCTAGGTGCCAAA	300
Db	GCAACCAAGATATCTATCTCTCTTTTGGCACCCATGAAACTGCATTTCTAGGTGCCAAA	787
QY	GACCTTTTCCATATAAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACAACCGAAAGGA	360
Db	GACCTTTTCCATATAAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACAACCGAAAGGA	847
QY	TTTAAACGAGGATGTGGGAAATAGAAAATAAACCAGGAGTAAAGTTTACTGGCTACCAG	420
Db	TTTAAACGAGGATGTGGGAAATAGAAAATAAACCAGGAGTAAAGTTTACTGGCTACCAG	907
QY	GCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAGTGAGAAATCTGCAGATGCA	480
Db	GCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAGTGAGAAATCTGCAGATGCA	967
QY	AGCAGTGAGGAAGAAGGTGATAGAGTAGAAGAAGATGAAAAGGCAAAAGAAAGAATGAA	540
Db	AGCAGTGAGGAAGAAGGTGATAGAGTAGAAGAAGATGAAAAGGCAAAAGAAAGAATGAA	1027
QY	AAAGCAGGCTCAAAACGGAAAAGTCTATCTTCAAGAAAATCTCTAAACAGTCCCGG	600
Db	AAAGCAGGCTCAAAACGGAAAAGTCTATCTTCAAGAAAATCTCTAAACAGTCCCGG	1087
QY	AAATCTCCAGGAGTGAAGATGACAAAGCTGCAAAGAGGAGAAACAAGACGAGCTCT	660
Db	AAATCTCCAGGAGTGAAGATGACAAAGCTGCAAAGAGGAGAAACAAGACGAGCTCT	1147
QY	GAGGGTGGAGATCGGGCAACGACACAAGAAAACAATCTTCAGATTCGAGAAAACCAAGT	720

Matches 1013; Conservative 1; Mismatches 3; Indels 4; Gaps 3;			
Qy	5	CTGTCGCGCGCGCTTGGAGCCGCGGGGAGCGC-GGGCAATTCGTGCGGCCGCGGGGG	63
Db	66	CTGTCGCGCGCGCTTGGAGCCGCGGGGAGCGGCAATTCGTGCGGCCGCGGGGG	125
Qy	64	GCGGCTCCCGGCATCTTCGGGGGCAACAAGACTACAGAAAGGGAGCGGTGGGATG	123
Db	126	GCGGCTCCCGGCATCTTCGGGGGCAACAAGACTACAGAAAGGGAGCGGTGGGATG	185
Qy	124	GCGGCTCCCGGGCCCGCGAGTACAAAGCGGCGACCTGCTTCGCCAAGATCAAGGC	183
Db	186	GCGGCTCCCGGGCCCGCGAGTACAAAGCGGCGACCTGCTTCGCCAAGATCAAGGC	245
Qy	184	TACCGCACTGGCGGCGCGGATGATGAATCCAGAGGCGCTGTGAAGCTCCAGCA	243
Db	246	TACCGCACTGGCGGCGCGGATGATGAATCCAGAGGCGCTGTGAAGCTCCAGCA	305
Qy	244	AACAGTATCTATCTCTCTTTTGGCACCCATGAATCTGATTTCTAGTCCCAAGAC	303
Db	306	AACAGTATCTATCTCTCTTTTGGCACCCATGAATCTGATTTCTAGTCCCAAGAC	365
Qy	304	CTTTTCCATATAGGAGTCAAGACAAGTTTGGAAAGTCMAACACGAAGGATTT	363
Db	366	CTTTTCCATATAGGAGTCAAGACAAGTTTGGAAAGTCMAACACGAAGGATTT	425
Qy	364	AACGAAGTGTGGGAAATAGAAATAACCCAGAGTAAAGTTTACTGGCTACCAAGCA	423
Db	426	AACGAAGTGTGGGAAATAGAAATAACCCAGAGTAAAGTTTACTGGCTACCAAGCA	485
Qy	424	ATTCAGCAACAGAGCTTTCAGAAACTGAGGGAGAGGTGGAATATCTGCAGATGCAAGC	483
Db	486	ATTCAGCAACAGAGCTTTCAGAAACTGAGGGAGAGGTGGAATATCTGCAGATGCAAGC	545
Qy	484	AGTGAGGAAGAGTGATAGTAGTAGAGAGATGGAAGGCAAAAGCAAGATGAAAA	543
Db	546	AGTGAGGAAGAGTGATAGTAGTAGAGATGGAAGGCAAAAGCAAGATGAAAA	605
Qy	544	GCAGGCTCAAAACGGGAAAGTATATATCTTCAAGAAATCTCTAAACAGTCCCGGAA	603
Db	606	GCAGGCTCAAAACGGGAAAGTATATATCTTCAAGAAATCTCTAAACAGTCCCGGAA	665
Qy	604	TCTCCAGGAGATGAAGATGACAAAGCTGCAAGAGAGAGGAAAAAACAACAGCTCTGAG	663
Db	666	TCTCCAGGAGATGAAGATGACAAAGCTGCAAGAGAGAGGAAAAAACAACAGCTCTGAG	725
Qy	664	GTTGGAGATCGGGCAACGACACAGAAACACAACTTCAGACTTGCAGAAAAACAGTAA	723
Db	726	GTTGGAGATCGGGCAACGACACAGAAACACAACTTCAGACTTGCAGAAAAACAGTAA	785
Qy	724	GGGACCTTAATACCATATGAATGCTGCATATTAAGAGAAACCAAGAAAGGTTATATGT	783
Db	786	GGGACCTTAATACCATATGAATGCTGCATATTAAGAGAAACCAAGAAAGGTTATATGT	845
Qy	784	TTGGTTGCTTAATATCTTGGATTTGATATGAACCAACACATAGTCTTGTGTCAATGA	843
Db	846	TTGGTTGCTTAATATCTTGGATTTGATATGAACCAACACATAGTCTTGTGTCAATGA	905
Qy	844	CAGAACCCAGTTTGATGTACATTAATTCATATTCCTCTCTGTGTGTTTCGGGGGAAA	903
Db	906	CAGAACCCAGTTTGATGTACATTAATTCATATTCCTCTCTGTGTGTTTCGGGGGAAA	965
Qy	904	AGACATTTAGCCCTTTTAAAGTTACTGATTAATTTTCATGTTATTTGTTGTCATGAA	963
Db	966	AGACATTTAGCCCTTTTAAAGTTACTGATTAATTTTCATGTTATTTGTTGTCATGAA	1025
Qy	964	GTTGCCCTTAACCACTAAGGATATCAAGATTTTGGCGAGACTTATACATGTCTAGGAT	1023
Db	1026	GTTGCC--TTAACACTAAGGATATCAAGATTTTGGCG--GACTTATACATGTCTAGGAT	1082
Qy	1024	C 1024	
Db	1083	C 1083	

RESULT 3
AL520469
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL520469 Homo sapiens 1201 bp mRNA linear EST 22-MAY-2003
AL520469 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
cDNA clone CS0DB006YG12 5-PRIME, mRNA sequence.
AL520469
AL520469.2 GI:31038810
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi:12783962.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7531.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DB006QPI&cluster=7531.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DB006BD06QPI.

FEATURES
Source

1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DB006YG12"
/issue_type="NEUROBLASTOMA COT 10-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 378 a 240 c 299 g 260 t 24 others

Query Match 88.6%; Score 907; DB 9; Length 1201; Best Local Similarity 98.9%; Pred. No. 9.9e-153; Matches 954; Conservative 1; Mismatches 6; Indels 4; Gaps 4;			
Qy	10	CGCGCCCGCTTGGAGCCCGCGGGAGCGC-GCGCAATTCGTGCGGCCGCGGGGGCGGC	68
Db	62	CGGGATGCTTGGAGCCCGCGGGAGCGGCGCAATTCGTGCGGCCGCGGGGGCGGC	121
Qy	69	CTCCCGGCATCTTTCGCGGCGACCAAGACTACACAGAAAGGGAGCGGCTGGGATGGCG	128
Db	122	CTCCCGGCATCTTTCGCGGCGACCAAGACTACACAGAAAGGGAGCGGCTGGGATGGCG	181
Qy	129	TGCGGGCCCGCGAGTACAAAGCGGGCGACCTGTGCTTCGCCAAGATGAGGGGTACCC	188
Db	182	TGCGGGCCCGCGAGTACAAAGCGGGCGACCTGTGCTTCGCCAAGATGAGGGGTACCC	241
Qy	189	GCACCTGGCGGGCCCGGATTTGATGAATCCACAGAGGCGCTGTGAAGCTCCACGAACAA	248
Db	242	GCACCTGGCGGGCCCGGATTTGATGAATCCACAGAGGCGCTGTGAAGCTCCACGAACAA	301
Qy	249	GTATCTCTATCTCTTTTGGCACCCATGAAACTGCAATTTCTAGGTGCCAAAGACCTTTT	308
Db	302	GTATCTCTATCTCTTTTGGCACCCATGAACTGCAATTTCTAGGTGCCAAAGACCTTTT	361
Qy	309	TCCATATAAGGAGTACAAAGACAGATTTTGGAAAGTCAAAACAAACGGAAGGATTTAACGA	368
Db	362	TCCATATAAGGAGTACAAAGACAGATTTTGGAAAGTCAAAACAAACGGAAGGATTTAACGA	421

```
QY 369 AGGATTGTGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACAGGCAATTCA 428
Db 422 AGGATTGTGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACAGGCAATTCA 481
QY 429 GCAACAGAGCTCTTCAGAACTGAGGAGAGGTGAAATACCTGAGATGCAAGCAGTGA 488
Db 482 GCAACAGAGCTCTTCAGAACTGAGGAGAGGTGAAATACCTGAGATGCAAGCAGTGA 541
QY 489 GGAAGAGGTGATAGAGTGAAGAGATGGAAGGCAAAAGAAAGAAAGTGAAGAGCAGG 548
Db 542 GGAAGAGGTGATAGAGTGAAGAGATGGAAGGCAAAAGAAAGAAAGTGAAGAGCAGG 601
QY 549 CTCAAAACGGAAAAGTCTATATCTTCAAGAAATCTCTTAACAGTCCCGAAATCTCC 608
Db 602 CTCAAAACGGAAAAGTCTATATCTTCAAGAAATCTCTTAACAGTCCCGAAATCTCC 661
QY 609 AGGAGATGAAGATGACAAAGACTGCAAGAGAGGAGAAACAAAGAGCAGCTCTGAGGGTGG 668
Db 662 AGGAGATGAAGATGACAAAGACTGCAAGAGAGGAGAAACAAAGAGCAGCTCTGAGGGTGG 721
QY 669 AGATGGGGCAAGCACAAGAAACACAACTTCAGACTTTGAGAGAAACCAAGTGAAGGAC 728
Db 722 AGATGGGGCAAGCACAAGAAACACAACTTCAGACTTTGAGAGAAACCAAGTGAAGGAC 781
QY 729 CTAACCTACCATATGAATGCTGCATATTAAGAGAAACCAAGAGGTTATATGTTGGT 788
Db 782 CTAACCTACCATATGAATGCTGCATATTAAGAGAAACCAAGAGGTTATATGTTGGT 841
QY 789 TGCTCAATATCTTGGAATTTGATATGAACCAACACATAGTCTCTGTTGTCATTGACAGAA 848
Db 842 TGCTCAATATCTTGGAATTTGATATGAACCAACACATAGTCTCTGTTGTCATTGACAGAA 901
QY 849 CCCAGTTTGTATGATATATTCATATTCCTCTCTGTTGTTGTCGGGGGAAAAGACA 908
Db 902 -CCCAGTTTGTATGATATATTCATATTCCTCTCTGTTGTTGTCGGGGGG-AAAAGACA 959
QY 909 TTTTAGCCTTTTAAAAGTACTGATTTAAATTTTCAATGTTATTTGTTGTCATGAAGTTGC 968
Db 960 TTTTAGCCTTTTAAAAG-TACTGATTTAAATTTTCAATGTTATTTGTTGTCATGAAGTKGC 1018
QY 969 CCTTA 973
Db 1019 CTAA 1023
```

```
RESULT 4
BG201738
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
```

BG201738 808 bp mRNA linear EST 21-APR-2001
RST21080 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
BG201738
BG201738.1 GI:13723425
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Harrington, J.J., Sherif, B., Rundlett, S., Jackson, P.D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J.,
Lerner, L., Costanzo, D., McElligott, K., Booser, S., Mays, R., Smith,
E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K., Offenbacher,
J., Danzig, J. and Ducar, M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
21227151
11329013
Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave., Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596

```
Email: scaine@thersys.com
High quality sequence stop: 511.
Location/Qualifiers
1..808
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/notes="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."
BASE COUNT 292 a 141 c 181 g 194 t
ORIGIN
Query Match 74.6%; Score 764; DB 10; Length 808;
Best Local Similarity 99.0%; Pred. No. 4.5e-127;
Matches 800; Conservative 0; Mismatches 5; Indels 3; Gaps 3;
QY 205 ATTGATGAATCCCGAGGGCGCTGTGAAGCTCCAGCAAAACAAGTATCTCTTCTTT 264
Db 1 ATTGATGAATCCCGAGGGCGCTGTGAAGCTCCAGCAAAACAAGTATCTCTTCTTT 60
QY 265 TTTGGCACCCATGAACCTGCAATTTCTAGGTCCTCAAGACCTTTTCCATATAAGGAGTAC 324
Db 61 TTTGGCACCCATGAACCTGCAATTTCTAGGTCCTCAAGACCTTTTCCATATAAGGAGTAC 120
QY 325 AAAGACAAGTTTGGAAAGTCAAAACAAACGAAAGGATTTAAACGAAGGATTTGTGGAAATA 384
Db 121 AAAGACAAGTTTGGAAAGTCAAAACAAACGAAAGGATTTAAACGAAGGATTTGTGGAAATA 180
QY 385 GAAATAATACCAGAGTAAAGTTTACTGGCTACAGGCAATTCAGCAACAGAGCTCTTCA 444
Db 181 GAAATAATACCAGAGTAAAGTTTACTGGCTACAGGCAATTCAGCAACAGAGCTCTTCA 240
QY 445 GAACTGAGGAGAGGTGGAATACTGCAGATCCAGCAGTGGAGGAGAGGTGATAGA 504
Db 241 GAACTGAGGAGAGGTGGAATACTGCAGATCCAGCAGTGGAGGAGAGGTGATAGA 300
QY 505 GTAGAAAGATGGAAGGCAAAAGAAAGAAATGAAAGAGGCTCAAAACGGAAAAAG 564
Db 301 GTAGAAAGATGGAAGGCAAAAGAAAGAAATGAAAGAGGCTCAAAACGGAAAAAG 360
QY 565 TCATATACCTCAAGAAATCTCTAAACAGTCCCGGAAATCTCCAGGAGATCAAGATGAC 624
Db 361 TCATATACCTCAAGAAATCTCTAAACAGTCCCGGAAATCTCCAGGAGATCAAGATGAC 420
QY 625 AAAGACTGCAAGAGAGGAAACAAAGAGCAGCTCTGAGGGTGGAGATCGCGGCAACGAC 684
Db 421 AAAGACTGCAAGAGAGGAAACAAAGAGCAGCTCTGAGGGTGGAGATCGCGGCAACGAC 480
QY 685 ACAAGAAACAACTTCAGACTTCAGAAAAACCAAGTGAAGGACCTAACTACCAATAATGA 744
Db 481 ACAAGAAACAACTTCAGACTTCAGAAAAACCAAGTGAAGGACCTAACTACCAATAATGA 540
QY 745 ATGCTGCATATTAAGAGAAACCAAGAGGTTATATCTTTGGTTGTCTAATATTTCTGG 804
Db 541 ATGCTGCATATTAAGAGAAACCAAGAGGTTATATCTTTGGTTGTCTAATATTTCTGG 600
QY 805 ATTTGATATGACCAACACATAGTCTTGTTCATTCAGACAGACCCAGTTGTATGTA 864
Db 601 ATTTGATATGACCAACACATAGTCTTGTTCATTCAGACAGACCCAGTTGTATGTA 660
QY 865 CATTATTTCATATTCCTCTCTGTTGTTTCCGGGGGAAAAGACATTTTAGCCCTTTTAA 924
Db 661 CATTATTTCATATTCCTCTCTGTTG-GGTTACAGAGGAAAGACATTTTAGCC-TTTTAA 718
QY 925 AAGTTACTGATTTAATTTTCATGTTTATTTGGTTGATGAAGTTGCCCTTAAACCACTAAGA 984
Db 719 AAGTTACTGATTTAATTTTCATGTTTATTTGGGTGCATGAAGTTGCCCTTTTACCACCTAAGGA 777
```

```
FEATURES
source
```

```

Qy 985 TTATCAAGATTTTTCGCGAGACTTATAC 1012
Db 778 TTATCAAGATTTTTCGCGAGACTTATAC 805

RESULT 5
BM457409 1112 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT 6407463 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5584879
DEFINITION 5', mRNA sequence.
ACCESSION BM457409
VERSION BM457409.1 GI:18506449
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1112)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12350 Row: e Column: 08
High quality sequence stop: 541.
FEATURES
source
1..1112
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5584879"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 385 a 211 c 228 g 288 t
ORIGIN
Query Match 74.5%; Score 763.2; DB 12; Length 1112;
Best Local Similarity 98.9%; Pred. No. 5.9e-127;
Matches 779; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Qy 236 CTCAGCAACAAGATCCTATCTCTTTTGGCACCCATGAACATGTCATTCTAGGTC 295
Db 1 CTCAGCAACAAGATCCTATCTCTTTTGGCACCCATGAACATGTCATTCTAGGTC 60

Qy 296 CCAAGACCTTTTCCATATAAGAGTACAAAGCAAGTTTGGAAAGTCAACAAACGGA 355
Db 61 CCAAGACCTTTTCCATATAAGAGTACAAAGCAAGTTTGGAAAGTCAACAAACGGA 120

Qy 356 AAGATTAAAGGAGGAGTTGGGAAATAGAAATACCCAGGAGTAAAGTTTACTGGCT 415
Db 121 AAGATTAAAGGAGGAGTTGGGAAATAGAAATACCCAGGAGTAAAGTTTACTGGCT 180

Qy 416 ACCAGGCAATTACGCAACAGAGCTCTTTAGAACTGAGGGAGAGGTGGAAATACTGCAG 475
Db 181 ACCAGGCAATTACGCAACAGAGCTCTTTAGAACTGAGGGAGAGGTGGAAATACTGCAG 240

Qy 476 ATGCAAGCAGTGAAGAGAGGTGATAGAGTAGAAGAGTGGAAAGGCAAGAAAGA 535
Db 241 ATGCAAGCAGTGAAGAGAGGTGATAGAGTAGAAGAGTGGAAAGGCAAGAAAGA 300

```

```

Qy 536 ATGAAAAAGCAGGCTCAAAACGGAAAAAGTCATATACTTCAAGAAATCCTCTAAACAGT 595
Db 301 ATGAAAAAGCAGGCTCAAAACGGAAAAAGTCATATACTTCAAGAAATCCTCTAAACAGT 360

Qy 596 CCCGGAATCTCCAGGAGATGAAGATGACAAAGCTGCAAGAGAGGAAAAACAAGCA 655
Db 361 CCCGGAATCTCCAGGAGATGAAGATGACAAAGCTGCAAGAGAGGAAAAACAAGCA 420

Qy 656 GCTCTGAGGGTGGAGATGCGGGCAACGACACAGAAACCAACTTCAGACTTCGACAGAAA 715
Db 421 GCTCTGAGGGTGGAGATGCGGGCAACGACACAGAAACCAACTTCAGACTTCGACAGAAA 480

Qy 716 CCAGTGAAGGGACCTAACTACCAATAATGAATGCTGCATATTTAAGAGAAAAACAAGAGG 775
Db 481 CCAGTGAAGGGACCTAACTACCAATAATGAATGCTGCATATTTAAGAGAAAAACAAGAGG 540

Qy 776 TTATATGTTGGTGTCTAATATCTTGGATTGATATGAAACCAACATAGTCCTTGT 835
Db 541 TTATATGTTGGTGTCTAATATCTTGGATTGATATGAAACCAACATAGTCCTTGT 600

Qy 836 GTCATTGACAGACCCAGTTTGTATGATACATTATTCATATTCCTCTCTGTCGTGTT-C 894
Db 601 GTCATTGACAGACCCAGTTTGTATGATACATTATTCATATTCCTCTCTGTCGTGTT-C 660

Qy 895 GGGGGGAAAAAGACATTTTAGCCCTTTTAAAGTTACTGATTTAATTTTCATGTTATTGG 954
Db 661 GGGGGGAAAAAGACATTTTAGCCCTTTTAAAGTTACTGATTTAATTTTCATGTTATTGG 720

Qy 955 TTGCATGAAGTTGCCCTTAAACCACTAAGGATTATCAAGATTTTGGCGAGACTTATACAT 1014
Db 721 TTGCATGAAGTTGCCCTTAAACCACTAAGGATTATCAAGATTTTGGCGAGACTTATACAT 780

Qy 1015 GTCCTAGGA 1022
Db 781 GTCCTAGGA 788

RESULT 6
AK044858 1868 bp mRNA linear HTC 05-DEC-2002
LOCUS Mus musculus 9.5 days embryo parthenogenote cDNA, RIKEN full-length
DEFINITION enriched library, clone:B130008G01 product:hepatoma-derived growth
factor, related protein 3, full insert sequence.
ACCESSION AK044858
VERSION AK044858.1 GI:26336878
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Fujimoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Yamawake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

```


RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, K.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M.,
Wagner, I., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
King, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
5

The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1668)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koyas, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Tissues were provided by Dr. Tomohiro Kono (Department of Animal
Science, Tokyo University of Agriculture, 1737 Hunko Atsugi City,
Kanagawa Prefecture, Japan) whose assistance we gratefully
acknowledge.

Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. .1868
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

/db_xref="FANTOM_DB:B130008G01"
 /db_xref="taxon:10090"
 /clone="B130008G01"
 /tissue_type="parthenogenote"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="9.5 days embryo"
 263..871
 /note="unnamed protein product; hepatoma-derived growth
 factor, related protein 3 (MGD) [MG:1352760, GB|NM_013886,
 evidence: BLASTN, 99%, match=611]
 putative"
 /codon_start=1
 /protein_id="BAC32123.1"
 /db_xref="GI:26336879"
 /translation="MARPPREYKAGDLVFAKMGYHPHARIDELPGAVKPPANKY
 PIFPFTHETAFGLPKDLPFYKEYKDFKGNKRRKFNWENNEPVFTGTQTI
 QQSSSTEGGNTADASSEEDRGVEDGKRNKKGSKRSKYSYTKSSKQSRK
 SPQSDDDKCKEENKSSSSEGGAGNDRNTTADLQAGEGT"
 BASE COUNT 519 a 414 c 457 g 478 t
 ORIGIN
 Query Match 74.0%; Score 758; DB 11; Length 1868;
 Best Local Similarity 86.5%; Pred. No. 4.5e-126;
 Matches 887; Conservative 0; Mismatches 125; Indels 14; Gaps 4;
 QY 2 CGGCTCGTCCGCCGGCTTGAGGCC---CGCGGGAGCGCGCGCAATTCTGTCGGCGCGG 58
 DB 141 CGGCTCGAGCGTTCGGCTCGAGGCGCGCGGGAGCGCGGCTTCGCTGTCGGCGG 200
 QY 59 GGGGGGGCGGCTCCCGGCGATCTTCGCGCGACCAAGAGACTACCAAGAAAGGGAGCGGCTG 118
 DB 201 GAGGGCGGCTCCGAGCATCTTCGCGCGCGCAAGAGCGCGGAGAGAGAACGCG 260
 QY 119 GGATGGCGGTCGCGCGGCCCGAGTACAAAGCGGGCGACCTGCTTTCGCGCAAGATGA 178
 DB 261 GGATGGCGGTCGCGCGGCCCGCGAGTACAAAGCGGGAGACCTGCTTTCGCGCAAGATGA 320
 QY 179 AGGGCTACCCGACTGCGCGCGCGGATGATGAATCTCCAGAGGGCGCTGTGAAGCCTC 238
 DB 321 AGGGCTACCCGACTGCGCGCGCGGATGATGAATCTCCAGAGGGAGCTGTGAAGCCTC 380
 QY 239 CAGCAAAACAAGTATCTATCTCTTTTGGCACCCATCAAACTGCAATTTCTAGTCCCA 298
 DB 381 CAGCAAAACAAGTATCTATCTCTTTTGGTACCCATGAACTGCAATTTCTAGTCTCTA 440
 QY 299 AAGACCTTTTCCATATAAGAGTACAAAGACAAGTTTGGAAAGTCAAAACCAACGGAAG 358
 DB 441 AAGACCTTTTCCATATAAGAGTACAAAGACAAGTTTGGAAAGTCAAAACCAACGGAAG 500
 QY 359 GATTTACGAGGATTTGGGGAATAGAAATACCCAGGAGTAAAGTTTACTGGCTACC 418
 DB 501 GATTTAATGAAGGATTTATGGGAAATATGAAATAATATCCAGGAGTGAATTTACTGGGTACC 560
 QY 419 AGGCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAGAGTGGAAATACTGCAGATG 478
 DB 561 AGCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAGAGTGGAAATACTGCAGATG 620
 QY 479 CAACAGTGAAGGAAGAAGGTGATAGTGAAGAGATGGAAGGCAAAAGAGAAAGATG 538
 DB 621 CAAGCAGTGAAGGAAGAAGGTGATAGT---AGAAGATGGAAGGCAAAAGAGAAAGATG 677
 QY 539 AAAAAGCAGGCTCAAAACGGGAAAGTCAATATCTTCAAGAAATCTCTTAACAGTCCC 598
 DB 678 AAAAAGGAGGCTCAAAACGGGAAAGTCTTACAGTCAAGAAAGTCTTCTTAACAGTCCC 737
 QY 599 GGAATCTCCAGGAGATCAAGATCAAGACTCAAAAGAGAGAGGAAACCAAAACAGCT 658
 DB 738 GGAATCTCCAGGAGAGCAGAGATGATTAAGACTTCNAAGAGAGAGGAGACAAACAGCT 797
 QY 659 CTGAGGGTGGAGATGCGGGCAACCGACACAGAAACACAACTTCAGACTTCGAGAAAAACA 718
 DB 798 CTGAGGGCGGAGACGCTGGCAATGACACGAGAAACACGACTTCGAGAAAGCGG 857

PUBMED
REFERENCE
AUTHORS11042159
3

Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS20530913
11076861
4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wegner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Togo-Oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS21085660
11217851
5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL
REFERENCE
AUTHORS

6 (bases 1 to 3363)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hangaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission

JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/

FEATURES
source

Location/Qualifiers

1. .3363
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="FANTOM DB:B230209J06"
/db_xref="taxon:10090"
/clone="B230209J06"
/sex="male"
/tissue_type="corpora quadrigemina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
438. .1100
CDS
/note="unnamed protein product; hepatoma-derived growth factor, related protein 3 (MGD|MG1:1352760, GB|NM_013886, evidence: BLASTN, 99%, match=611)
putative"

CDS

BASE COUNT 981 a 670 c 793 g 919 t
ORIGIN

Query Match 72.9%; Score 746; DB 11; Length 3363;
Best Local Similarity 86.4%; Pred. No. 5.6e-124;
Matches 886; Conservative 0; Mismatches 125; Indels 15; Gaps 5;
QY 2 CCGCTCGTCCGCGCGCTTCAGGCC---CGCGGGAGCGCGCAATTCGTGCGCGCGCG 58
Db 316 CCGCTCGAGGTTTCGGCTCGAGGCGCGCGCGCGCGGTTTCGTGCTGCGCGCG 375
QY 59 GGGGGGCGGCTCCCGGCATCTTCGGCGGCACCAAGGACTACCAAGAGGGGAGCGGCTG 118
Db 376 GAGGGGCGGCTCCAGCATCTTCGGCGCGCGCAAGGACCGCGGAGCAGAGAACGACG 435
QY 119 GATGCGGCTCGCGCGCGCGGAGTACAAAGCGGCGACCTGCTTTCGCCAAGATGA 178
Db 436 GATGCGGCTCGCGCGCGCGGAGTACAAAGCGGAGACCTGCTTTCGCCAAGATGA 495
QY 179 AGGCTACCGCGACTGCGCGCGCGGATTCGATGAATCCAGAGGGCGCTGTGAAGCTC 238
Db 496 AGGCTACCGCGACTGCGCGCGCGGATTCGATGAATCCAGAGGGAGCTGTGAAGCTC 555
QY 239 CAGCAACAAAGTATCTTCTCTTTTGGCACCCATGAATCTGCAATTTCTAGGTCCCA 298
Db 556 CAGCAACAAAGTATCTTCTCTTTTGGTACCCATGAATCTGCAATTTCTAGGTCTTA 615
QY 299 AAGACCTTTTTCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAG 358
Db 616 AAGACCTTTTCCCATATAGGAATACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAG 675
QY 359 GATTTAAGGAGGATTTGGGAATAGNAATTAACCCAGGAGTAAAGTTTACTGGCTACC 418
Db 676 GATTTAAGGAGGATTTGGGAATAGNAATTAACCCAGGAGTAAAGTTTACTGGGTACC 735
QY 419 AGGCAATTCAGCAACAGAGCTCTTCAGAACTCAGGAGAGAGGTGGAAATCTATGCGAG 478
Db 736 AGCAATTCAGCAACAGAGCTCTTCAGAACTCAGGAGAGAGGTGGAAATCTATGCGAG 795
QY 479 CAAGCAGTAGGAGAGGTTGATAGAGTAGAGTAGAGAGAGAGAGAGAGAGAGAGAGAG 538
Db 796 CAAGCAGTAGGAGAGGTTGATAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAG 852
QY 539 AAAAGCAGGCTCAAAACGGAAGTATATATCTATCAAGAAATCTCTTAAACAGTCCC 598
Db 853 AAAAGCAGGCTCAAAACGGAAGTATATATCTATCAAGAAATCTCTTAAACAGTCCC 912
QY 599 GGAAATCTCCAGGAGATGAAGATGACAAAGACTGCAAGAGAGAGAGAGAGAGAGAGAG 658

Db 913 GGAATCT-CAGGAGACGAGATGATAAAGACTGCAAGAGAGGAGAAACAAAGCAGCT 971
 QY 659 CTGAGGGTGGAGATCGGCAACGACACAGAAACACAACTTCAGACTTGCAGAAACCA 718
 Db 972 CTGAGGGCGGAGACGCTGGCAATGACACGAGAAACACGACTGAGACTTGCAGAAAGCG 1031
 QY 719 GTGAAGGACCTAACTACCAATATGCTGATTAAGAGAAACCAACAGAGGTTA 778
 Db 1032 GTGAAGGACCTAACTACCGTAATGCTGCAATTCAGAGACACCAAGAGGTTA 1091
 QY 779 TAGTTTGGTGTCTAATATCTTGGATTGATGAACCAACATAGTCTCTTGTGTC 838
 Db 1092 ACTGTTTGTATGCTAATCTTGGATTGATGAACCAACAGGCTCTTGTGTC 1151
 QY 839 ATTGACAGAACCCAGTTGTATGATATATTCATATTCCTCTGTTGTGTTCCGGG 898
 Db 1152 ATTGACAGAACCCAGTATGTTGTCATTTAGTATGAACCAACAGGCTCTTGTGTC 1211
 QY 899 GGAAGAGACATTTAGCTTTTAAAGTTACTGATTTAAATTCATGTTATTTGTTGC 958
 Db 1212 AAAGAGACATTTAAACCTTTAA-----GGTGTATTTATTTATTTGTTGC 1266
 QY 959 ATGAAGTCCCTTAAACCACTAAGGATTAATCAAGATTTTGGCAGACTTATACATGCT 1018
 Db 1267 ATGAAGTCCCTTAACTAAGATGATGAAGATTTT---GCAGACTTCTGTGTCT 1323
 QY 1019 AGGATC 1024
 Db 1324 AGGCTC 1329

RESULT 11
 BG204817
 LOCUS
 DEFINITION
 RST24232 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
 ACCESSION
 BG204817
 VERSION
 BG204817.1 GI:13726504
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)

ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 819)
 HARRINGTON, J.J., SHERF, B., RUNDLETT, S., JACKSON, P.D., PERRY, R.,
 CAIN, S., LEVENTHAL, C., THORNTON, M., RAMACHANDRAN, R., WHITTINGTON, J.,
 LERNER, L., COSTANZO, D., MCELLIGOTT, K., BOOZER, S., MAYERS, R., SMITH,
 J., VELOSO, N., KLIKA, A., HESS, J., COthren, K., LO, K., OFFENBACHER,
 J., DANZIG, J. and DUCAR, M.
 Creation of genome-wide protein expression libraries using random
 activation of gene expression

TITLE
 Creation of genome-wide protein expression libraries using random
 activation of gene expression

JOURNAL
 Nat. Biotechnol. 19 (5), 440-445 (2001)
 MEDLINE
 21227151
 PUBMED
 11329013
 COMMENT
 Contact: Scott J. Cain
 Athersys, Inc.
 3201 Carnegie Ave., Cleveland, OH 44115, USA
 Tel: 216 431 9900
 Fax: 216 361 9596
 Email: scain@athersys.com
 High quality sequence stop: 551.

FEATURES
 Location/Qualifiers
 1..819
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="HT1080"
 /clone_lib="Athersys RAGE Library"
 /note="See 'Creation of Genome-wide Protein Expression
 Libraries using Random Activation of Gene Expression',
 Nature Biotechnology, in press. Note that even though the
 cell type indicated is HT1080, since a random activation
 method was used, these sequence tags are not necessarily
 EST."

BASE COUNT 289 a 149 c 180 g 199 t expressed in HT1080 under normal circumstances."
 ORIGIN
 Query Match 72.2%; Score 739; DB 10; Length 819;
 Best Local Similarity 98.2%; Pred. No. 1.4e-122;
 Matches 778; Conservative 0; Mismatches 11; Indels 3; Gaps 3;
 QY 196 CCGGCCCGGATTGATGAACCTCCAGAGGGCGCTGTGAAGCCCTCCAGCAAAACAAGTATCT 255
 Db 31 CCGCTACAGATTGATGAACCTCCAGAGGGCGCTGTGAAGCCCTCCAGCAAAACAAGTATCT 90
 QY 256 ATCTCTTTTGGCACCCCATGAACTGCAATTTCTAGTCCCAAGACCTTTTCCATAT 315
 Db 91 ATCTCTTTTGGCACCCCATGAACTGCAATTTCTAGTCCCAAGACCTTTTCCATAT 150
 QY 316 AAGGAGTACAAAGACAGATTTTGGAAAGTCAAAACAAACGGAAGGATTTAAACGAAGATTG 375
 Db 151 AAGGAGTACAAAGACAGATTTTGGAAAGTCAAAACAAACGGAAGGATTTAAACGAAGATTG 209
 QY 376 TGGGAAATAGAAAATAAACCAGGAGTAAAGTTTACTGCTACCGGCAATTCAGCAACAG 435
 Db 210 TGGGAAATAGAAAATAAACCAGTAAATTTTACTGCTACCGGCAATTCAGCAACAG 269
 QY 436 AGCTCTTCAGAACTGAGGAGAGGTGGAAATACTGCAGATGCAAGCAGTGAGGAAGAA 495
 Db 270 AGCTCTTCAGAACTGAGGAGAGGTGGAAATACTGCAGATGCAAGCAGTGAGGAAGAA 329
 QY 496 GGTGATAGTAGAAGAGATGGAAGGCAAAAGCAAAAGAGATGAAAAGCAGGCTCAAAA 555
 Db 330 GGTGATAGTAGAAGAGATGGAAGGCAAAAGCAAAAGAGATGAAAAGCAGGCTCAAAA 389
 QY 556 CGGAAAAAGTCAATATCTTCAAGAAATCCTCTTAAACAGTCCCGGAAATCTCCAGAGAT 615
 Db 330 CGGAAAAAGTCAATATCTTCAAGAAATCCTCTTAAACAGTCCCGGAAATCTCCAGAGAT 449
 QY 616 GAAAGTACAAAGACTGCAAAAGAGAGGAAACCAAAAGCAGCTCTGAGGGTGGAGATGCG 675
 Db 450 GAAAGTACAAAGACTGCAAAAGAGAGGAAACCAAAAGCAGCTCTGAGGGTGGAGATGCG 509
 QY 676 GGCAACGACACAGAAACACAACTTCAGACTTGCAGAAACCAAGTGAAGGGACCTAACTA 735
 Db 510 GGCAACGACACAGAAACACAACTTCAGACTTGCAGAAACCAAGTGAAGGGACCTAACTA 569
 QY 736 CCATATGAATGCTGCATATTAAGAGAAACCAACAGAGAGGTATATGTTTGGTCTCTAA 795
 Db 570 CCATATGAATGCTGCATATTAAGAGAAACCAACAGAGAGGTATATGTTTGGTCTCTAA 629
 QY 796 TATCTTGGATTGATGAACCAACACATAGTCTCTTGTGTCATTTGACAGAACCCCACT 855
 Db 630 TATCTTGGATTGATGAACCAACACATAGTCTCTTGTGTCATTTGACAGAACCCCACT 689
 QY 856 TTGATGATACATTTATTCATTTCTCTCTGTTGTTGTTTGGGGGGAAGAGACATTTTACG 915
 Db 690 TTGATGATACATTTATTCATTTCTCTCTGTTGTTGTTTGGGGGGAAGAGACATTTTACG 748
 QY 916 CTTTCTTAAAGTTTACTGATTTAAATTTTCATGTTATTTGTTGTCATGAAGATTCGCCCTTAC 975
 Db 749 C-TTTTAAAGTTTACTGATTTAAATTTTCATGTTATTTGTTGTCATGAAGATTCGCCCTTAC 807
 QY 976 CACTAAGGATTA 987
 Db 808 CACTAAGGATTA 819

RESULT 12
 BM722135
 LOCUS
 DEFINITION
 UI-E-E00-ahy-c-17-0-UI.r1 UI-E-E00 Homo sapiens cDNA clone
 UI-E-E00-ahy-c-17-0-UI 5', mRNA sequence.
 ACCESSION
 BM722135
 VERSION
 BM722135.1 GI:19042502
 KEYWORDS
 EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 692)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBR, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hegeman
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers
 1..692
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-800-ahy-c-17-0-UI"
 /tissue_type="fetal eye"
 /dev_stage="fetal"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-E-800"
 /notes="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 UI-E-800 is a cDNA library containing the following
 tissue(s): fetal eye. The library was constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pT73-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dr)18 tail. The
 sequence tag for this library is CGGTATACC. This library
 was created for the program, Gene Discovery in the Visual
 System, supported by National Eye Institute (NEI)."
 BASE COUNT 257 a 111 c 157 g 167 t

ORIGIN

Query Match 67.4%; Score 690.4; DB 12; Length 692;
 Best Local Similarity 99.9%; Pred. No. 7.2e-114;
 Matches 691; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 329 ACAAGTTTGCAAGTCAAAACGGAAGGATTTTACGAGGATTTGCGGAATAGAAA 388
 DB 1 ACAAGTTTGGAAGTCAAAACGGAAGGATTTTACGAGGATTTGCGGAATAGAAA 60
 QY 389 ATACCCAGGAGTAAAGTTTACTGGCTTACCGGCAATTCAGCAACAGAGCTCTTCAGAAA 448
 DB 61 ATACCCAGGAGTAAAGTTTACTGGCTTACCGGCAATTCAGCAACAGAGCTCTTCAGAAA 120
 QY 449 CTGAGGAGAGGTGGAAATCTGCAGATCAACAGCTGAGGAAGAGGTGATAGATAG 508
 DB 121 CTGAGGAGAGGTGGAAATCTGCAGATCAACAGCTGAGGAAGAGGTGATAGATAG 180
 QY 509 ACAAGATGCAAAAGCGCAAAAGAAAGATGAAAAGCAGAGCTCAAAACGGAAAAAGTCAT 568
 DB 181 AAGAAGATGCAAAAGCGCAAAAGAAAGATGAAAAGCAGAGCTCAAAACGGAAAAAGTCAT 240

QY 569 ATACTTTCAAAGAAATCCTCTAAACAGTCCCGGAAATCTCCAGGAGATGAAGATGACAAAG 628
 DB 241 ATACTTTCAAAGAAATCCTCTAAACAGTCCCGGAAATCTCCAGGAGATGAAGATGACAAAG 300
 QY 629 ACTGCAAGAGAGGAGAAACAAAGACAGCTCTGAGGGTGGAGATGCGGGCAACGACACAA 688
 DB 301 ACTGCAAGAGAGGAGAAACAAAGACAGCTCTGAGGGTGGAGATGCGGGCAACGACACAA 360
 QY 689 GAAACACAACTTCAGACTTGCAGAAAAACAGTGAAGGACCTAACTACCAATAATGAATGC 748
 DB 361 GAAACACAACTTCAGACTTGCAGAAAAACAGTGAAGGACCTAACTACCAATAATGAATGC 420
 QY 749 TGCATATTAAAGAGAAACCAAGAGGTTATATGTTGGTTGTCTAATATTCTTGGATTT 808
 DB 421 TGCATATTAAAGAGAAACCAAGAGGTTATATGTTGGTTGTCTAATATTCTTGGATTT 480
 QY 809 GATATGAACCAACACATAGTCTTGTGTCATTGACAGAACCCAGTTTGTATGTACATT 868
 DB 481 GATATGAACCAACACATAGTCTTGTGTCATTGACAGAACCCAGTTTGTATGTACATT 540
 QY 869 ATTCAATATTCCTCTCTGTTGTGTTTCGGGGGAAAAAGACATTTTAGCCTTTTTTAAAAAGT 928
 DB 541 ATTCAATATTCCTCTCTGTTGTGTTTCGGGGGAAAAAGACATTTTAGCCTTTTTTAAAAAGT 600
 QY 929 TACTGATTTAAATTTCAATGTTATTTGTTGCATGAAGTTGCCCTTAACCACTAAGGATTAT 988
 DB 601 TACTGATTTAAATTTCAATGTTATTTGTTGCATGAAGTTGCCCTTAACCACTAAGGATTAT 660
 QY 989 CAAGATTTTGGCAGACATTATACATGCTCTAG 1020
 DB 661 CAAGATTTTGGCAGACATTATACATGCTCTAG 692

RESULT 13
 BE886964 781 bp mRNA linear EST 20-OCT-2000
 LOCUS 601506447F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3908171 5',
 DEFINITION mRNA sequence.
 ACCESSION BE886964
 VERSION BE886964.1 GI:10341775
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 781)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM9719 row: n column: 12
 High quality sequence stop: 766.
 Location/Qualifiers
 1..781
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3908171"
 /issue_type="leiomysarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_71"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
 Average insert size 2.1 kb."
 BASE COUNT 285 a 129 c 168 g 199 t

ORIGIN	Query Match	65.6%; Score 672.2; DB 10; Length 781;
	Best Local Similarity	99.4%; Pred. No. 1.3e-110; Length 781;
	Matches 685; Conservative	0; Mismatches 3; Indels 1; Gaps 1;
QY	336	TGGAAGTCAACAAACGGAAGGATTTAAACGAAGGATTTGTGGGAAATAGAAAATAAACCC 395
DB	1	TGGAAGTCAACAAACGGAAGGATTTAAACGAAGGATTTGTGGGAAATAGAAAATAAACCC 60
QY	396	AGGAGTAAGTTTACTGGCTACAGGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGG 455
DB	61	AGGAGTAAGTTTACTGGCTACAGGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGG 120
QY	456	AGAAGGTGAATATCTGCAGATGCAAGCAGTGCAGGAGAGGTGATAGTAGAGAAGA 515
DB	121	AGAAGGTGAATATCTGCAGATGCAAGCAGTGCAGGAGAGGTGATAGTAGAGAAGA 180
QY	516	TGGAAGGCAAAAGAAAGATGAAAAAGCAGGCTCAAAACGGAAGGATCATATCTTC 575
DB	181	TGGAAGGCAAAAGAAAGATGAAAAAGCAGGCTCAAAACGGAAGGATCATATCTTC 240
QY	576	AAAGAAATCTCTAAACAGTCCCGAAATCTCCAGAGATGAAGATGCAAGAGCTGCAA 635
DB	241	AAAGAAATCTCTAAACAGTCCCGAAATCTCCAGAGATGAAGATGCAAGAGCTGCAA 300
QY	636	AGAAGGGAACAAACAGAGCTCTGAGGCTGGAGTGGGCAACGACACAAAGAACAC 695
DB	301	AGAAGGGAACAAACAGAGCTCTGAGGCTGGAGTGGGCAACGACACAAAGAACAC 360
QY	696	AACTTCAGACTTCGAGAAACAGTGAAGGACCTAACTTACCATAATGAATGCTCATAT 755
DB	361	AACTTCAGACTTCGAGAAACAGTGAAGGACCTAACTTACCATAATGAATGCTCATAT 420
QY	756	TAAAGAAACCAAGAGGTTATATGTTGGTGTCTGAATATCTTGGATTTGATAGA 815
DB	421	TAAAGAAACCAAGAGGTTATATGTTGGTGTCTGAATATCTTGGATTTGATAGA 480
QY	816	ACCAACATAGTCTTGTGTCATGACAGAACCCAGTTTGTATGTATTCATATTCATA 875
DB	481	ACCAACATAGTCTTGTGTCATGACAGAACCCAGTTTGTATGTATTCATATTCATA 540
QY	876	TTCTCTCTGTGTTTCGGGGGAAAGACATTTTAGCCCTTTTAAAAAGTTACTGAT 935
DB	541	TTCTCTCTGTGTTTCGGGGGAAAGACATTTTAGCCCTTTTAAAAAGTTACTGAT 599
QY	936	TTAATTTCAATGTTATTTGGTTCATGAAGTTGCCCTTAACCACTAAGGATTAAGATT 995
DB	600	TTAATTTCAATGTTATTTGGTTCATGAAGTTGCCCTTAACCACTAAGGATTAAGATT 659
QY	996	TTTGGCGAGCTTATACATGCTAGGATC 1024
DB	660	TTTGGCGAGCTTATACATGCTAGGATC 688
RESULT 14		
LOCUS	BG212737	
DEFINITION	BG212737	813 bp mRNA linear EST 21-APR-2001
ACCESSION	BG212737	RS23232 Athersys RAGE Library Homo sapiens cdNA, mRNA sequence.
VERSION	BG212737.1	GI:13734424
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE		
AUTHORS	Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R., Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S., Mays, R., Smith, E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K., Offenbacher, J., Danzig, J. and Ducar, M.	
TITLE	Creation of genome-wide protein expression libraries using random	

JOURNAL	activation of gene expression
MEDLINE	Nat. Biotechnol. 19 (5), 440-445 (2001)
PUBMED	21227151
COMMENT	11329013
	Contact: Scott J. Cain
	Athersys, Inc.
	3201 Carnegie Ave, Cleveland, OH 44115, USA
	Tel: 216 431 9900
	Fax: 216 361 9596
	Email: scain@atersys.com
FEATURES	High quality sequence stop: 380.
source	Location/Qualifiers
	1..813
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/cell_line="HT1080"
	/clone_lib="Athersys RAGE Library"
	/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."
BASE COUNT	295 a 138 c 197 g 183 t
ORIGIN	
Query Match	64.4%; Score 659.6; DB 10; Length 813;
Best Local Similarity	93.5%; Pred. No. 2.3e-108;
Matches 754; Conservative	0; Mismatches 44; Indels 8; Gaps 6;
QY	205 ATTGATGAATCCCAAGAGGGCGCTGTGAAGCTCCAGCAAAACAAGTATCTCTTCTTT 264
Db	1 ATTGATGAATCCCAAGAGGGCGCTGTGAAGCTCCAGCAAAACAAGTATCTCTTCTTT 60
QY	265 TTTGGCACCCATGAACTGCAATTTCTAGGTCCTCCAAAGACCTTTTCCATATAAGGAGTAC 324
Db	61 TTTGGCACCCATGAACTGCAATTTCTAGGTCCTCCAAAGACCTTTTCCATATAAGGAGTAC 120
QY	325 AAAGCAAGTTTGGAAAGTCAACAAACGGAAGGATTTAAGGAGGATTTGGGGAATAA 384
Db	121 AAAGCAAGTTTGGAAAGTCAACAAACGGAAGGATTTAAGGAGGATTTGGGGAATAA 180
QY	385 GAAATAATACCCAGAGTAAAGTTTACTGGCTACAGGCAATTCAGCAACAGAGCTCTTCA 444
Db	181 GAAATAATACCCAGAGTAAAGTTTACTGGCTACAGGCAATTCAGCAACAGAGCTCTTCA 240
QY	445 GAAACTGAGGAGAGGTGGAAATATCTGCAGATGCAAGCAGTGAGGAGAGAGTGATAGA 504
Db	241 GAAACTGAGGAGAGGTGGAAATATCTGCAGATGCAAGCAGTGAGGAGAGAGTGATAGA 300
QY	505 GTAGAAGAGATGGAAGGCAAAAGAAAGAAAGTGAAGGAGGCTCAAAACCGGAAAAAG 564
Db	301 GTAGAAGAGATGGAAGGCAAAAGAAAGAAAGTGAAGGAGGCTCAAAACCGGAAAAAG 360
QY	565 TCATATATCTTCAAGAAATCTCTTAAACAGTCCCGGAAATCTCCAGGAGATGAAGATGAC 624
Db	361 TCATATATCTTCAAGAAATCTCTTAAACAGTCCCGGAAATCTCCAGGAGATGAAGATGAC 420
QY	625 AAAGACTGCAAGAGGAGGAAACAAACAGCAGCTCTGAGGGTGGAGATGCGGGCAACGAC 684
Db	421 AAAGACTGCAAGAGGAGGAAACAAACAGCAGCTCTGAGGGTGGAGATGCGGGCAACGAC 480
QY	685 ACAGAAACACACTTCAGACTTCGAGAAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAG 744
Db	481 ACAGAAACACACTTCAGACTTCGAGAAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAG 540
QY	745 ATGCTGCATATTAAGAG-AAACCAACAAGAGTTTATAT-GTTTGGTGTGCTAATATCTTT 802
Db	541 ATGCTGCATATTAAGAGAAACCAAGAGAGTTTATATGTTTGGTGGTGGTGGTGGTGGTGG 600
QY	803 GGATTGTGATGAACCAACACATAGTCTTGTGTCATTGACAGAACCCCAAGTTTGTATG 862

FT Modified-site 149 /note= "Potential cAMP- and cGMP-dependent protein kinase phosphorylation site"

FT Modified-site 151 /note= "Potential protein kinase C phosphorylation site"

FT Modified-site 152 /note= "Potential protein kinase C phosphorylation site"

FT Modified-site 155 /note= "Potential protein kinase C phosphorylation site"

FT Modified-site 156 /note= "Potential cAMP- and cGMP-dependent protein kinase phosphorylation site"

FT Modified-site 159 /note= "Potential protein kinase C phosphorylation site"

FT Modified-site 162 /note= "Potential casein kinase II phosphorylation site"

FT Modified-site 176 /note= "Potential N-glycosylation site"

FT Modified-site 178 /note= "Potential casein kinase II phosphorylation site"

FT Modified-site 187 /note= "Potential N-glycosylation site"

FT Modified-site 191 /note= "Potential N-glycosylation site"

FT Modified-site 192 /note= "Potential casein kinase II phosphorylation site"

XX WO9398976-A2.

XX 05-AUG-1999.

XX 12-JAN-1999; 99WO-US000654.

XX 29-JAN-1998; 98US-0015412.

XX (INCY-) INCYTE PHARM INC.

XX Baughn M, Corley NC, Guegler KJ, Hillman JL, Lal P; Tang YT;

XX WPI; 1999-469331/39.

XX N-PSDB; AA200048.

XX New human growth factor homologs and their use in the diagnosis, treatment and prevention of cell proliferative and developmental disorders

XX Claim 1; Fig 2; 89pp; English.

XX This sequence is the HGF2 Human Growth Factor Homologue 2. The HGF2 protein has structural and chemical homology to the human hepatoma derived growth factor (HGF). HGF2 is 203 amino acids in length and has three potential N-glycosylation sites, two potential cAMP- and cGMP-dependent protein kinase phosphorylation sites, eight potential casein kinase II phosphorylation sites and six potential protein kinase C phosphorylation sites. HGF2s are expressed in cancerous tissues and appear to play a role in cell proliferative and developmental disorders. A purified antagonist of HGF can be administered to a subject as a method of treating or preventing a cell proliferative disorder e.g. atherosclerosis, hepatitis, and some cancers. Antibodies against HGF and antisense sequences may also be used as antagonist for treating the above. Developmental disorders treated by HGF include renal tubular acidosis, anaemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, and epilepsy. HGF antibodies and HGF polynucleotides and polypeptides may also be used in various diagnostic methods.

XX Sequence 203 AA;

SQ Query Match 100.0%; Score 1086; DB 20; Length 203;

Best Local Similarity 100.0%; Pred. No. 2.5e-97;

Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARPRPREYKAGDLVFAKMGYPHWPARIDELPEGAVKPPANKYPPIFFGTHETAFLGPK 60

DB 1 MARPRPREYKAGDLVFAKMGYPHWPARIDELPEGAVKPPANKYPPIFFGTHETAFLGPK 60

QY 61 DLFPYKEYKDFGKSNKRKGNFENGLWEIENNPVGKFTGYQAIQQSSSETEGEGNTADA 120

DB 61 DLFPYKEYKDFGKSNKRKGNFENGLWEIENNPVGKFTGYQAIQQSSSETEGEGNTADA 120

QY 121 SSEEGRDVRVEDGKGRKNEKAGSKRKSYTSKKSKOSRKS PGDEDDKCKEENKSSS 180

DB 121 SSEEGRDVRVEDGKGRKNEKAGSKRKSYTSKKSKOSRKS PGDEDDKCKEENKSSS 180

QY 181 EGGDAGNDTRNTTSDLOKTSEGT 203

DB 181 EGGDAGNDTRNTTSDLOKTSEGT 203

RESULT 2

AABI3521

ID AABI3521 standard; Protein; 203 AA.

XX AABI3521;

AC AABI3521;

XX 13-NOV-2000 (first entry)

DT Human hepatoma-derived growth factor homologous polypeptide.

XX Human; hepatoma-derived growth factor homologous polypeptide; huHGFh; cell proliferation; haematopoiesis; angiogenesis;

KW lymphocyte proliferation; infection; autoimmune disease; vascular disease; cancer.

XX Homo sapiens.

OS WO200037492-A2.

PN 29-JUN-2000.

PD 22-DEC-1999; 99WO-US30932.

PF 22-DEC-1999; 98US-0113344.

PR (ELIL) LILLY & CO ELI.

XX Na S;

XX WPI; 2000-442638/38.

DR N-PSDB; AAA65004.

XX Novel human homologues of hepatoma-derived growth factor homologous nucleic acids, polypeptides useful as probes or amplification primers in the detection, quantitation or isolation of gene sequences or transcripts

XX Claim 9; Page 74-75; 75pp; English.

XX The present sequence is the human homologue of the hepatoma-derived growth factor (huHGFh). This protein is involved in tumour formation in some cells, and it is thought that the gene and protein will be useful in the diagnosis and treatment of infections, autoimmune disorders, vascular diseases and cancers. In addition, the gene can be used to produce transgenic animals which may be used as animal models for these diseases, and the protein can be used to induce cell proliferation, haematopoiesis, lymphocyte proliferation and angiogenesis.

XX Sequence 203 AA;

SQ Query Match 100.0%; Score 1086; DB 21; Length 203;

Best Local Similarity 100.0%; Pred. No. 2.5e-97;

Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARPRPREYKAGDLVFAKMGYPHWPARIDELPEGAVKPPANKYPPIFFGTHETAFLGPK 60

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2004, 15:26:51 ; Search time 41 Seconds
(without alignments)
785.890 Million cell updates/sec

Title: US-09-787-328B-4

Perfect score: 1086

Sequence: 1 MARPRPYKAGDLVFAKMK.....DAGNDTRNTTSLDQKTSST 203

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1086	100.0	203	AA228502	HGFH2 Human Growth
2	1086	100.0	203	AA228502	Human hepatoma-der
3	1086	100.0	203	AA228502	Human type II hepa
4	1086	100.0	203	AA228502	Human protein sequ
5	1086	100.0	203	AA228502	Human secreted pro
6	1086	100.0	203	AA228502	Human albumin fusi
7	1053.5	97.0	202	AA228502	Rat protein isolat
8	965	88.9	244	AA228502	Novel human diagno
9	940	86.6	203	AA228502	Human hepatoma-der

10	895	82.4	172	21	AA228502	Human cancer assoc
11	672.5	61.9	197	22	ABG15332	Novel human diagno
12	535.5	49.3	246	21	AA228502	Lung cancer associ
13	534.5	49.2	240	16	AA228502	Human hepatoma der
14	534.5	49.2	240	21	AA228502	Human HDGFIP prote
15	534.5	49.2	240	22	AB250263	Hepatoma-derived g
16	532.5	49.0	235	19	AA228502	Lung growth factor
17	532	49.0	237	18	AA228502	Mouse hepatoma der
18	525	48.3	326	22	AB221116	Human PRO1604 homo
19	524	48.3	314	22	AB228502	Human secreted pro
20	524	48.3	314	23	AB228502	Human albumin fusi
21	524	48.3	671	21	AA228502	Human PRO1604 (UNQ
22	524	48.3	671	22	AA228502	Human PRO polypt
23	524	48.3	671	22	AA228502	Protein of the inv
24	524	48.3	671	23	AA228502	Human LP protein L
25	524	48.3	671	23	AB228502	Human NOV4d protei
26	524	48.3	671	24	AB228502	Human PRO1604 prot
27	524	48.3	671	24	AB228502	Human secreted/tra
28	524	48.3	671	24	AB228502	Novel human secret
29	524	48.3	671	24	AB228502	Human secreted/tra
30	524	48.3	671	24	AB228502	Human PRO polypt
31	524	48.3	671	24	AB228502	Human PRO polypt
32	524	48.3	671	24	AB228502	Human secreted/tra
33	524	48.3	671	24	AB228502	Human secreted/tra
34	524	48.3	671	24	AB228502	Human secreted/tra
35	524	48.3	676	22	AA228502	Human hepatome cel
36	523	48.2	676	22	AA228502	Human NOV9 protein
37	522	48.1	667	23	AB228502	Human NOV4b protei
38	522	48.1	667	23	AB228502	Human NOV4c protei
39	504.5	46.5	669	19	AA228502	Mouse liver cancer
40	504	46.4	676	22	AA228502	Human 'SB-HDGF prot
41	503	46.3	676	23	AB228502	Human NOV4a protei
42	486.5	44.8	530	20	AA228502	Human lens epithel
43	486.5	44.8	530	23	AA228502	Human lens epithel
44	485.5	44.7	325	21	AA228502	Human p75 protein
45	485.5	44.7	333	21	AA228502	Human p52 protein.

ALIGNMENTS

RESULT 1

AA228502

ID AA228502 standard; Protein; 203 AA.

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

Db 1 MARPPREYKAGDLVFAKMGYPHWPARIIDELPEGAVKPPANKYPIFFGTHETAFLGPK 60
Qy 61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGKFTGYQAIQQSSSETEGGGNTADA 120
Db 61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGKFTGYQAIQQSSSETEGGGNTADA 120
Qy 121 SSEEGRVVEDGKGRKNEKAGSKRKYSYTSKKSSKQSRKSPGDEDDKCKEENKSSS 180
Db 121 SSEEGRVVEDGKGRKNEKAGSKRKYSYTSKKSSKQSRKSPGDEDDKCKEENKSSS 180
Qy 181 EGGDAGNDRNTTSDLOKTSEGT 203
Db 181 EGGDAGNDRNTTSDLOKTSEGT 203

RESULT 3
AA98374
ID AAY88374 standard; Protein; 203 AA.
XX
AC AAY88374;
DT 20-JUL-2000 (first entry)
XX
DE Human type II hepatoma derived growth factor (HDGF2) amino acid sequence.
XX
KW Hepatoma-derived growth factor 2; HDGF2; cytostatic; drug;
XX treatment; hepatoma.
OS Homo sapiens.
XX
PN W0200017351-A1.
XX
PD 30-MAR-2000.
XX
PF 06-SEP-1999; 99WO-CN00139.
XX
PR 22-SEP-1998; 98CN-0119758.
XX
PA (YULL/) YU L.
XX
PI Yu L, Zhang H, Fu Q, Zhao Y, Tu Q;
XX WPI; 2000-283579/24.
DR N-PSDB; AAA13161.
XX
PS Type II human-derived growth factor with homology to its type I
PT version, useful e.g. in study and development of drugs for hepatoma
XX
PS Claim 2; Page 15; 27pp; Chinese.
XX
CC This sequence represents the human hepatoma-derived growth factor (HDGF2)
CC protein sequence. The HDGF2 sequence shows considerable homology to the
CC type I version. The protein has cytostatic activity. The HDGF2 protein,
CC derivatives and polynucleotides are useful e.g. in the study and
CC development of drugs for the treatment of hepatoma.
XX
SQ Sequence 203 AA;

Query Match 100.0%; Score 1086; DB 21; Length 203;
Best Local Similarity 100.0%; Pred. No. 2.5e-97;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARPPREYKAGDLVFAKMGYPHWPARIIDELPEGAVKPPANKYPIFFGTHETAFLGPK 60
Db 1 MARPPREYKAGDLVFAKMGYPHWPARIIDELPEGAVKPPANKYPIFFGTHETAFLGPK 60
Qy 61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGKFTGYQAIQQSSSETEGGGNTADA 120
Db 61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGKFTGYQAIQQSSSETEGGGNTADA 120
Qy 121 SSEEGRVVEDGKGRKNEKAGSKRKYSYTSKKSSKQSRKSPGDEDDKCKEENKSSS 180
Db 121 SSEEGRVVEDGKGRKNEKAGSKRKYSYTSKKSSKQSRKSPGDEDDKCKEENKSSS 180

Qy 181 EGGDAGNDRNTTSDLOKTSEGT 203
Db 181 EGGDAGNDRNTTSDLOKTSEGT 203

RESULT 4
AA92800
ID AAB92800 standard; Protein; 203 AA.
XX
AC AAB92800;
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:11307.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
XX
PR 27-AUG-1999; 99JP-0300253.
XX
PR 11-JAN-2000; 2000JP-0118776.
XX
PR 02-MAY-2000; 2000JP-0183767.
XX
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 11307; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AA03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AA092446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 203 AA;

Query Match 100.0%; Score 1086; DB 22; Length 203;
Best Local Similarity 100.0%; Pred. No. 2.5e-97;

Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARPPREYKAGDLVFAKMGYPHPARIDELPEGAVPPANKYPIFFFGTHETAFLGPK 60
 DB 1 MARPPREYKAGDLVFAKMGYPHPARIDELPEGAVPPANKYPIFFFGTHETAFLGPK 60
 QY 61 DLFPYKEYKDKFGKSKRKGKFNGLWEIENNPVGFYQAIQQOSSSETEGEGNTADA 120
 DB 61 DLFPYKEYKDKFGKSKRKGKFNGLWEIENNPVGFYQAIQQOSSSETEGEGNTADA 120
 QY 121 SSEEGRVVEEDGKGRKNEKAGSKRKKSYTSKSSKQSRKSPGDEDDKCKEENKSSS 180
 DB 121 SSEEGRVVEEDGKGRKNEKAGSKRKKSYTSKSSKQSRKSPGDEDDKCKEENKSSS 180
 QY 181 EGGDAGNDRNTTSDLOKTSEGT 203
 DB 181 EGGDAGNDRNTTSDLOKTSEGT 203
 RESULT 5
 AAB70073
 ID AAB70073 standard; Protein; 203 AA.
 AC AAB70073;
 DT 14-MAY-2001 (first entry)
 XX Human secreted protein #12.
 DE Human; secreted protein; immunomodulatory; anti-sclerotic;
 KW dermatological; immunosuppressive; anti-inflammatory; anti-HIV;
 KW immunostimulant; cytostatic; cardiant; vascular; anti-angiogenic;
 KW ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;
 KW antialzheimers; antiparkinsonian; antimicrobial; vulnary; gene therapy;
 KW immune disorder; hyperproliferative; cardiovascular; angiogenic;
 KW neurological; infection.
 XX Homo sapiens.
 OS
 KW WO200112776-A2.
 PN
 XX
 PD 22-FEB-2001.
 XX
 PF 15-AUG-2000; 2000WO-US22350.
 PR 16-AUG-1999; 99US-0148759.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Shi Y, Young PE, Ebner R, Soppet DR, Ruben SM;
 PI
 XX WPI; 2001-244245/25.
 DR N-PSDB; AAF76854.
 XX
 XX Nucleic acids encoding 18 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
 PT disease and diabetic retinopathy -
 XX
 XX Claim 11; Page 369; 380pp; English.
 CC
 XX The present sequence is one of 18 novel human secreted proteins. The
 CC nucleic acids encoding the proteins and the proteins themselves may be
 CC used in the prevention, diagnosis and treatment of diseases including
 CC immune disorders (e.g. multiple sclerosis, systemic lupus erythematosus
 CC and human immunodeficiency virus (HIV) infections), hyperproliferative
 CC disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases
 CC (e.g. Scimitar syndrome, Chaga's cardiomyopathy and coronary
 CC arteriosclerosis), angiogenic disorders (e.g. corneal graft
 CC neovascularisation and diabetic retinopathy), neurological disorders
 CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),
 CC infectious diseases and/or for promoting wound healing, regeneration
 CC and/or chemotaxis. The nucleic acid molecules may also be used as DNA
 CC probes in diagnostic assays to detect and quantitate the presence of

CC similar nucleic acid sequences in samples. The polypeptides may also be
 CC used as antigens in the production of antibodies and in assays to
 CC identify modulators of protein expression and activity.
 XX
 SQ Sequence 203 AA;
 Query Match 100.0%; Score 1086; DB 22; Length 203;
 Best Local Similarity 100.0%; Pred. No. 2.5e-97;
 Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARPPREYKAGDLVFAKMGYPHPARIDELPEGAVPPANKYPIFFFGTHETAFLGPK 60
 DB 1 MARPPREYKAGDLVFAKMGYPHPARIDELPEGAVPPANKYPIFFFGTHETAFLGPK 60
 QY 61 DLFPYKEYKDKFGKSKRKGKFNGLWEIENNPVGFYQAIQQOSSSETEGEGNTADA 120
 DB 61 DLFPYKEYKDKFGKSKRKGKFNGLWEIENNPVGFYQAIQQOSSSETEGEGNTADA 120
 QY 121 SSEEGRVVEEDGKGRKNEKAGSKRKKSYTSKSSKQSRKSPGDEDDKCKEENKSSS 180
 DB 121 SSEEGRVVEEDGKGRKNEKAGSKRKKSYTSKSSKQSRKSPGDEDDKCKEENKSSS 180
 QY 181 EGGDAGNDRNTTSDLOKTSEGT 203
 DB 181 EGGDAGNDRNTTSDLOKTSEGT 203
 RESULT 6
 AABG65515
 ID AABG65515 standard; Protein; 203 AA.
 AC AABG65515;
 XX
 DT 27-AUG-2002 (first entry)
 XX Human albumin fusion protein #2190.
 DE
 XX Albumin fusion protein; therapeutic protein X; human albumin; HA;
 KW human serum albumin; HSA; cancer; reproductive disorder;
 KW digestive disorder; immune disorder; endocrine disorder;
 KW haematopoietic disorder; neural disorder; connective disorder;
 KW cytostatic; antiinfertility; antiinflammatory; antitumor;
 KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
 KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
 KW osteopathic; antiarthritic.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200177137-A1.
 XX
 PD 18-OCT-2001.
 XX
 PF 12-APR-2001; 2001WO-US11988.
 XX
 PR 12-APR-2000; 2000US-229358P.
 PR 25-APR-2000; 2000US-199384P.
 PR 21-DEC-2000; 2000US-256931P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Haseltine WA;
 PI
 XX WPI; 2002-010886/01.
 DR
 XX New fusion protein for treating disease e.g. diabetes comprises an
 PT albumin fused to a therapeutic protein -
 PT
 XX Claim 1; Page 2083-2084; 2102pp; English.
 PS
 XX The present invention relates to albumin fusion proteins comprising a
 CC therapeutic protein X and human albumin (HA), also known as human serum
 CC albumin, HSA). The proteins are useful for treating a disease or

CC disorder that may be modulated by therapeutic protein X. The albumin
 CC extends the shelf-life of protein X, and may increase its biological
 CC in vitro/in vivo activity. The protein is useful for treating and
 CC diagnosing disorders such as cancer, reproductive disorders, digestive
 CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
 CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
 CC (e.g. diabetes), hematopoietic disorders, neural disorders
 CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
 CC encephalomyelitis, meningitis, schizophrenia), and connective disorders
 CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin
 CC fusion proteins of the invention.
 XX
 SQ Sequence 203 AA;

Query Match 100.0%; Score 1086; DB 23; Length 203;
 Best Local Similarity 100.0%; Pred. No. 2.5e-97;
 Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MARPPREYKAGDLVFAKMGYPHPARIDELPEGAVKPPANKYPIFFFGTHETAFILGPK 60
 DB 1 MARPPREYKAGDLVFAKMGYPHPARIDELPEGAVKPPANKYPIFFFGTHETAFILGPK 60
 QY 61 DLPPYKEYKDKFGKSKRKGFEGLWEIENNPVGKFTGYQAIQQSSSETEGGNTADA 120
 DB 61 DLPPYKEYKDKFGKSKRKGFEGLWEIENNPVGKFTGYQAIQQSSSETEGGNTADA 120
 QY 121 SSEEGRVVEDGKGRKNEKAGSKRKSYTSKSKSKSRKSPGDEDDKCKEENKSSS 180
 DB 121 SSEEGRVVEDGKGRKNEKAGSKRKSYTSKSKSKSRKSPGDEDDKCKEENKSSS 180
 QY 181 EGGDAGNDTRNTTSDLQKTSEGT 203
 DB 181 EGGDAGNDTRNTTSDLQKTSEGT 203

RESULT 7
 ABB72327
 ID ABB72327 standard; Protein; 202 AA.
 XX
 AC ABB72327;
 XX
 DT 04-APR-2002 (first entry)
 XX
 DE Rat protein isolated from skin cells SEQ ID NO: 651.
 XX
 KW Human; rat; mouse; skin call; skin wound; cancer; growth defect;
 KW developmental defect; inflammatory disease; dermatological; vulnary;
 KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.
 XX
 OS Rattus sp.
 XX
 PN WO200190357-A1.
 XX
 PD 29-NOV-2001.
 XX
 PF 24-MAY-2001; 2001WO-NZ00099.
 XX
 PR 24-MAY-2000; 2000US-206650P.
 PR 25-JUL-2000; 2000US-221232P.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX
 PI Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG, Kumble KD;
 PI WPI; 2002-122020/16.
 DR N-PSDB; ABLJ5012.
 DR
 XX New polynucleotides and polypeptides encoded by the polynucleotides
 PT isolated from skin cells, useful for treating skin wounds, cancers,
 PT growth and developmental defects, inflammatory diseases, or for
 PT modulating immune responses -
 XX
 PS Claim 4; Page 408-409; 466pp; English.

XX The present invention provides the protein and coding sequences of cDNAs
 CC isolated from human, murine and rat skin cell libraries. The sequences
 CC can be used in the development of therapeutic agents useful in the
 CC treatment of skin diseases, including skin wounds, cancer, growth
 CC defects, developmental defects and inflammatory diseases. The proteins
 CC have important roles in the induction of hair growth, cell proliferation
 CC and cell-cell interaction, in maintaining tissue integrity, in wound
 CC healing and in modulating immune responses. The present sequence is a
 CC polypeptide of the invention.
 XX
 SQ Sequence 202 AA;

Query Match 97.0%; Score 1053.5; DB 23; Length 202;
 Best Local Similarity 97.5%; Pred. No. 3.5e-94;
 Matches 198; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
 QY 1 MARPPREYKAGDLVFAKMGYPHPARIDELPEGAVKPPANKYPIFFFGTHETAFILGPK 60
 DB 1 MARPPREYKAGDLVFAKMGYPHPARIDELPEGAVKPPANKYPIFFFGTHETAFILGPK 60
 QY 61 DLPPYKEYKDKFGKSKRKGFEGLWEIENNPVGKFTGYQAIQQSSSETEGGNTADA 120
 DB 61 DLPPYKEYKDKFGKSKRKGFEGLWEIENNPVGKFTGYQAIQQSSSETEGGNTADA 120
 QY 121 SSEEGRVVEDGKGRKNEKAGSKRKSYTSKSKSKSRKSPGDEDDKCKEENKSSS 180
 DB 121 SSEEGRVVEDGKGRKNEKAGSKRKSYTSKSKSKSRKSPGDEDDKCKEENKSSS 180
 QY 181 EGGDAGNDTRNTTSDLQKTSEGT 203
 DB 180 EGGDAGNDTRNTTSDLQKAGEGT 202

RESULT 8
 ABG15333
 ID ABG15333 standard; Protein; 244 AA.
 XX
 AC ABG15333;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #15324.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 PI WPI; 2001-639362/73.
 DR N-PSDB; AAS79520.
 DR
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID No 45692; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and

CC involved in tumour formation in some cells, and it is thought that the
 CC gene and protein will be useful in the diagnosis and treatment of
 CC infections, autoimmune disorders, vascular diseases and cancers. In
 CC addition, the gene can be used to produce transgenic animals which may be
 CC used as animal models for these diseases, and the protein can be used to
 CC induce cell proliferation, haematopoiesis, lymphocyte proliferation and
 CC angiogenesis.
 CC Note: This sequence is not shown in the specification, but is derived
 CC from the wild-type human homologue of the hepatoma-derived growth factor
 CC shown in SEQ ID NO: 2 (AAB13521).

XX
 SQ Sequence 203 AA;
 Query Match 86.6%; Score 940; DB 21; Length 203;
 Best Local Similarity 85.2%; Pred. No. 3.7e-83;
 Matches 173; Conservative 11; Mismatches 19; Indels 0; Gaps 0;
 QY 1 MARPRREYKAGDLVFAKMGKYPHPARIDLPAGAVKPPANKYPIFFFGTHETAFILGPK 60
 DB 1 MARPRREFPGDILFAKMGKYPHPARVDDXADGAVKPPXNKLPIFFFGTHETAFILGPK 60
 QY 61 DLPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGKFTGYQAIQQOSSSETEGEGNTADA 120
 DB 61 DIFPYXXKXKYGKPNKRKGFNEGLWEIENNPXASYSXXPPIXXSSSETEGEGNTADA 120
 QY 121 SSEEGDRVEDGKGRKNEKAGSKKSYTSKSSKQSRKSPGDEDDKCKEENKSSS 180
 DB 121 SSEEGDRVEDGKGRKNEKAGSKKSYTSKSSKQSRKSPGDEDDKCKEENKSSS 180
 QY 181 EGGDAGNDTRNTTSDLOKTSBGT 203
 DB 181 EGGDAGNDTRNTTSDLOKTSBGT 203

RESULT 10
 AAB43542
 ID AAB43542 standard; Protein; 172 AA.
 XX
 AC AAB43542;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human cancer associated protein sequence SEQ ID NO:987.
 XX
 KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnerable; immunomodulator;
 KW antidiabetic; antisthmatic; antirheumatic; antithyroid; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.
 XX
 OS Homo sapiens.
 XX
 FN WO200055350-A1.
 XX
 PD 21-SEP-2000.
 XX
 XX 08-MAR-2000; 2000WO-US05882.
 XX
 PF 12-MAR-1999; 99US-0124270.
 XX
 PR (HUMA-) HUMAN GENOME SCI INC.
 XX
 PA Rosen CA, Ruben SM;
 XX
 PI WPI; 2000-587533/55.
 XX
 DR N-PSDB; AAC77751.
 XX
 PT Novel isolated nucleic acids comprising sequences encoding peptides

PT useful for treating or diagnosing e.g. cancer -
 XX Claim 11; Page 1562; 2352pp; English.
 XX
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given
 CC in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnerable; immunomodulator;
 CC antidiabetic; antisthmatic; antirheumatic; antithyroid; antiviral;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; coagulant;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.

XX Sequence 172 AA;
 SQ
 Query Match 82.4%; Score 895; DB 21; Length 172;
 Best Local Similarity 98.8%; Pred. No. 6.9e-79;
 Matches 169; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 33 PEGAVKPPANKYPIFFFGTHETAFILGPKDLFPYKEYKDKFGKSNKRKGFNEGLWEIENNP 92
 DB 2 PRGAVKPSANKYPIFFFGTHETAFILGPKDLFPYKEYKDKFGKSNKRKGFNEGLWEIENNP 61
 QY 93 GVKFTGYQAIQQOSSSETEGEGNTADASSEEGDRVEDGKGRKNEKAGSKKSYTS 152
 DB 62 GVKFTGYQAIQQOSSSETEGEGNTADASSEEGDRVEDGKGRKNEKAGSKKSYTS 121
 QY 153 KXSSKQSRKSPGDEDDKCKEENKSSSEGGDAGNDTRNTTSDLOKTSBGT 203
 DB 122 KXSSKQSRKSPGDEDDKCKEENKSSSEGGDAGNDTRNTTSDLOKTSBGT 172

RESULT 11
 ABG15332
 ID ABG15332 standard; Protein; 197 AA.
 XX
 AC ABG15332;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #15323.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 FN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 XX
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.
 DR N-PSDB; AAS79519.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID No 45691; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AEG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: the sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 197 AA;
 Query Match 61.9%; Score 672.5; DB 22; Length 197;
 Best Local Similarity 87.2%; Pred. No. 3.3e-57;
 Matches 129; Conservative 3; Mismatches 7; Indels 9; Gaps 1;
 QY 19 MKGYHPHWPAR-----IDELPGAVKPPANKYPIFFFGTHETAFGLPKDLFPYKEYK 69
 Db 35 LRAAPLQPTDRPSDMVAFIDELPGAVKPPANKYPIFFFGTHETAFGLPKDLFPYKEYK 94
 QY 70 DFGKSNKRGKGFNEGLWEIENNPVGKFTGYQAIQQSSSTETEGEGNTADASSEEGDRV 129
 Db 95 DFGKSNKRGKGFNEGLWEIENNPVGKFTGYQAIQQSSSTETEGEGNTADASSEEGDRV 154
 QY 130 EDGKGRKNEKAGSKRKXGYSYTSKSSK 157
 Db 155 EDGKGRKNEKAGSKRKXGYSYTSKTK 182
 RESULT 12
 ID AAB58173
 AC AAB58173 standard; Protein; 246 AA.
 XX
 AC AAB58173;
 XX
 DT 14-MAR-2001 (first entry)
 XX
 DE Lung cancer associated polypeptide sequence SEQ ID 511.
 XX
 KW Human; lung cancer associated protein; neuroprotective; cytostatic;
 KW cardioactive; immunomodulatory; muscular active; vulnary;
 KW gastrointestinal; nephrotropic; antiinfective; gynecological;
 KW antibacterial; diagnosis; neural disorder; immune disorder;
 KW proliferative disorder; wound healing; infectious disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200055180-A2.
 XX
 PD 21-SEP-2000.
 XX

PF 08-MAR-2000; 2000WO-US05918.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 XX
 PI Ruben SM;
 XX
 DR WPI; 2000-587514/55.
 DR N-PSDB; AAF18049.
 XX
 PT Lung cancer associated gene sequences, referred to as lung cancer
 PT antigens, useful for treatment, prevention, and diagnosis of disorders
 PT such as lung cancer -
 XX
 PS Claim 11; Page 999-1000; 1425pp; English.
 XX
 CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
 CC associated proteins and polynucleotide sequences, their agonists, and
 CC antagonists may have neuroprotective; cytostatic; cardioactive;
 CC immunomodulatory; muscular active general; vulnary; gastrointestinal
 CC general; nephrotropic; antiinfective; gynecological; or antibacterial
 CC activity. The invention also includes antibodies specific for the
 CC protein or polynucleotide sequences. The lung cancer associated
 CC polynucleotide sequences may be used for detection of lung cancer,
 CC chromosome identification, as chromosome markers, and for numerous other
 CC diagnostic or research purposes. The proteins may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders. The proteins may also be used in the treatment of wounds and
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
 CC peptide AAB58549 are used in the course of the invention for the
 CC identification and characterisation of the polynucleotide and protein
 CC sequences.
 XX
 SQ Sequence 246 AA;
 Query Match 49.3%; Score 535.5; DB 21; Length 246;
 Best Local Similarity 51.4%; Pred. No. 8.7e-44;
 Matches 114; Conservative 25; Mismatches 58; Indels 25; Gaps 6;
 QY 5 RPREYKAGDLVFAKMKGYHPHWPARIDELPGAVKPPANKYPIFFFGTHETAFGLPKDLFP 64
 Db 12 RQKEYKCGDLVFAKMKGYHPHWPARIDENPEAAVKSTANKYQVFFGTHETAFGLPKDLFP 71
 QY 65 YKEYKDKFGKSNKRGKGFNEGLWEIENNPVGKFTGYQAIQQSSSSP-----TEGEG-- 114
 Db 72 YEESEKFKGKPNKRGKGFSEGLWEIENNPVTKASGYQSSQKSKSCVPEPEPEAAEGDGD 131
 QY 115 -GNTADASSEEGDRVVEEDGKGRKNEKAGSKRKKS---YTSKSSKSKRSPGDEDDKD 170
 Db 132 KGN-AEGSSDBEG-KLVIDEPAKNEKNGALKRAGDILLEDSPKPKPAENPEGEKEAA 189
 QY 171 CKE-----EENKSSSEGGDAGNDTRNTTSDLOKTSSEGT 203
 Db 190 TLEVERPLPMEVEKNSTXSEPGSGRGPQEEEEEDEEEAT 231
 RESULT 13
 AAR66727
 ID AAR66727 standard; Protein; 240 AA.
 XX
 AC AAR66727;
 XX
 DT 02-OCT-1995 (first entry)
 XX
 DE Human hepatoma derived growth factor.
 XX
 KW Human hepatoma derived growth factor; recombinant production;
 KW receptor purification.
 XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2004, 15:26:53 ; Search time 22 seconds
(without alignments)
390.414 Million cell updates/sec

Title: US-09-787-328B-4

Perfect score: 1086

Sequence: 1 MARPPREYKAGDLVFAKMK.....DAGNDRNTTSDIQTSEGT 203

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	534.5	49.2	240	US-08-760-745-3	Sequence 3, Appli
2	532.5	49.0	235	US-08-760-745-1	Sequence 1, Appli
3	532	49.0	237	US-08-760-745-5	Sequence 5, Appli
4	133.5	12.3	764	US-09-370-838-67	Sequence 67, Appl
5	119.5	11.0	302	US-09-282-305-14	Sequence 14, Appl
6	119.5	11.0	302	US-09-883-720-14	Sequence 14, Appl
7	118	10.9	305	US-09-282-305-12	Sequence 12, Appl
8	118	10.9	305	US-09-883-720-12	Sequence 12, Appl
9	117.5	10.8	414	US-07-667-276A-4	Sequence 4, Appli
10	114	10.5	445	US-08-845-258-38	Sequence 38, Appl
11	114	10.5	445	US-08-950-571-38	Sequence 38, Appl
12	114	10.5	445	US-08-723-142A-38	Sequence 38, Appl
13	114	10.5	445	US-09-528-784A-38	Sequence 38, Appl
14	114	10.5	445	US-09-569-098A-38	Sequence 38, Appl
15	114	10.5	666	US-09-528-784A-85	Sequence 85, Appl
16	114	10.5	666	US-09-569-098A-85	Sequence 85, Appl
17	114	10.5	1132	US-09-528-784A-87	Sequence 87, Appl
18	114	10.5	1132	US-09-569-098A-87	Sequence 87, Appl
19	112	10.3	231	US-09-461-697-194	Sequence 194, App
20	112	10.3	232	US-09-461-697-192	Sequence 192, App
21	112	10.3	238	US-09-461-697-190	Sequence 190, App
22	112	10.3	257	US-09-461-697-188	Sequence 188, App
23	112	10.3	272	US-09-461-697-186	Sequence 186, App
24	112	10.3	311	US-09-282-305-16	Sequence 16, Appl
25	112	10.3	311	US-09-883-720-16	Sequence 16, Appl
26	111.5	10.3	482	US-08-528-255A-1	Sequence 1, Appli
27	111.5	10.3	482	US-08-717-365-1	Sequence 1, Appli

28	109.5	10.1	415	3	US-08-675-816-2	Sequence 2, Appli
29	107.5	9.9	1085	1	US-08-431-080-28	Sequence 28, Appl
30	107.5	9.9	1085	2	US-08-938-534-28	Sequence 28, Appl
31	107.5	9.9	1085	4	US-09-345-294-28	Sequence 28, Appl
32	105.5	9.7	720	4	US-09-480-921B-8	Sequence 8, Appli
33	105.5	9.7	1850	4	US-09-620-093A-5	Sequence 5, Appli
34	102.5	9.4	913	4	US-08-971-089-4	Sequence 4, Appli
35	100	9.2	531	2	US-08-933-750C-9	Sequence 9, Appli
36	100	9.2	531	3	US-09-234-613-9	Sequence 9, Appli
37	99	9.1	210	4	US-09-252-991A-19969	Sequence 19969, A
38	98	9.0	317	6	5340934-11	Patent No. 5340934
39	98	9.0	754	3	US-09-214-564A-2	Sequence 2, Appli
40	97.5	9.0	1162	2	US-08-728-323A-2	Sequence 2, Appli
41	97.5	9.0	1162	4	US-09-298-568-2	Sequence 2, Appli
42	97.5	9.0	1162	4	US-09-410-399-2	Sequence 2, Appli
43	97	8.9	182	1	US-08-466-603-5	Sequence 5, Appli
44	97	8.9	182	1	US-08-314-503A-5	Sequence 5, Appli
45	97	8.9	182	1	US-08-468-066-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-760-745-3
; Sequence 3, Application US/08760745
; Patent No. 5972658
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, surya K.
; APPLICANT: Murty, Lytn E.
; TITLE OF INVENTION: LUNG GROWTH FACTOR VARIANT
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,745
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0169 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 598956
; US-08-760-745-3

Query Match 49.2%; Score 534.5; DB 2; Length 240;
Best Local Similarity 51.4%; Pred. No. 1.6e-46;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2004, 15:29:12 ; Search time 33 Seconds
(without alignments)
1257.869 Million cell updates/sec

Title: US-09-787-328B-4
Perfect score: 1086
Sequence: 1 MARPPREYKAGDLVFAKMK.....DAGNDTRNTSDLQKTSEGT 203

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 762491 seqs, 204481190 residues

Total number of hits satisfying chosen parameters: 762491

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : Published Applications AA.*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1086	100.0	203	9	US-09-788-826-41
2	1086	100.0	203	12	US-10-247-671-132
3	1033.5	97.0	202	11	US-09-866-050A-651
4	895	82.4	172	9	US-09-925-301-987
5	749	69.0	145	9	US-09-768-826-60
6	535.5	49.3	246	9	US-09-925-302-511
7	534.5	49.2	240	9	US-09-938-885A-3
8	534.5	49.2	240	11	US-09-987-755-7
9	534.5	49.2	240	12	US-10-116-275-220
10	534.5	49.2	240	15	US-10-207-791-2
11	532.5	49.0	235	9	US-09-938-885A-1
12	532	49.0	237	9	US-09-938-885A-5
13	524	48.3	314	9	US-09-768-826-43
14	524	48.3	670	11	US-09-823-187-86
15	524	48.3	670	12	US-09-863-776-51

16	524	48.3	671	11	US-09-946-374-308	Sequence 308, App
17	524	48.3	671	11	US-09-823-187-85	Sequence 85, Appl
18	524	48.3	671	12	US-10-015-387A-308	Sequence 308, App
19	524	48.3	671	12	US-10-006-130A-308	Sequence 308, App
20	524	48.3	671	12	US-10-199-672-346	Sequence 346, App
21	524	48.3	671	12	US-10-006-172A-308	Sequence 308, App
22	524	48.3	671	12	US-10-187-749-346	Sequence 346, App
23	524	48.3	671	12	US-10-194-457-346	Sequence 346, App
24	524	48.3	671	12	US-10-184-642-346	Sequence 346, App
25	524	48.3	671	12	US-10-196-747-346	Sequence 346, App
26	524	48.3	671	12	US-10-015-392A-308	Sequence 308, App
27	524	48.3	671	12	US-10-017-253A-308	Sequence 308, App
28	524	48.3	671	12	US-10-173-689-346	Sequence 346, App
29	524	48.3	671	12	US-10-173-690-346	Sequence 346, App
30	524	48.3	671	12	US-10-173-691-346	Sequence 346, App
31	524	48.3	671	12	US-10-173-692-346	Sequence 346, App
32	524	48.3	671	12	US-10-173-694-346	Sequence 346, App
33	524	48.3	671	12	US-10-173-698-346	Sequence 346, App
34	524	48.3	671	12	US-10-173-709-346	Sequence 346, App
35	524	48.3	671	12	US-10-173-707-346	Sequence 346, App
36	524	48.3	671	12	US-10-174-569-346	Sequence 346, App
37	524	48.3	671	12	US-10-174-583-346	Sequence 346, App
38	524	48.3	671	12	US-10-174-587-346	Sequence 346, App
39	524	48.3	671	12	US-10-174-589-346	Sequence 346, App
40	524	48.3	671	12	US-10-174-591-346	Sequence 346, App
41	524	48.3	671	12	US-10-175-736-346	Sequence 346, App
42	524	48.3	671	12	US-10-175-742-346	Sequence 346, App
43	524	48.3	671	12	US-10-175-744-346	Sequence 346, App
44	524	48.3	671	12	US-10-175-745-346	Sequence 346, App
45	524	48.3	671	12	US-10-175-748-346	Sequence 346, App

ALIGNMENTS

RESULT 1
US-09-768-826-41
; Sequence 41, Application US/09768826
; Patent No. US20020012966A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: 18 human secreted proteins
; FILE REFERENCE: PF512P1
; CURRENT APPLICATION NUMBER: US/09768, 826
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/US00/22350
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/148, 759
; PRIOR FILING DATE: 1999-08-16
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-768-826-41

Query Match	100.0%	Score 1086;	DB 9;	Length 203;
Best Local Similarity	100.0%	Pred. No. 4.9e-82;		
Matches 203;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MARPPREYKAGDLVFAKMKGYPHWPARIDELPEGAVKPPANKYPPIFFGTHETAFGLCPK	60	
Db	1	MARPPREYKAGDLVFAKMKGYPHWPARIDELPEGAVKPPANKYPPIFFGTHETAFGLCPK	60	
Qy	61	DLFPYKEYKDKFGSKNKRKGFNEGLWEIENNPVGKFTGYQAIQQSSSETEGEGGNTADA	120	
Db	61	DLFPYKEYKDKFGSKNKRKGFNEGLWEIENNPVGKFTGYQAIQQSSSETEGEGGNTADA	120	
Qy	121	SSEEGDRVDEGDKGRKNEKAGSKRKYSYTSKKSKQSKRSGPDEDDDKCKKEENKSSS	180	
Db	121	SSEEGDRVDEGDKGRKNEKAGSKRKYSYTSKKSKQSKRSGPDEDDDKCKKEENKSSS	180	

```
QY 181 EGGDAGNDTRNTTSDLOKTSSEGT 203
Db 181 EGGDAGNDTRNTTSDLOKTSSEGT 203

RESULT 2
US-10-247-671-132
; Sequence 132, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 132
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 2676869CD1
US-10-247-671-132

Query Match 100.0%; Score 1086; DB 12; Length 203;
Best Local Similarity 100.0%; Pred. No. 4.9e-82; Indels 0; Gaps 0;
Matches 203; Conservative 0; Mismatches 0;

QY 1 MARPPREYKAGDLVFAKMKGYPHWPARIDELPEGAVKPPANKYPIFFFGTHETAFGLGPK 60
Db 1 MARPPREYKAGDLVFAKMKGYPHWPARIDELPEGAVKPPANKYPIFFFGTHETAFGLGPK 60

QY 61 DLFPYKEYKDKFGKSNKRGFNEGLWEIENNPVGVKFTGYQAIQQSSSSETEGEGNTADA 120
Db 61 DLFPYKEYKDKFGKSNKRGFNEGLWEIENNPVGVKFTGYQAIQQSSSSETEGEGNTADA 120

QY 121 SSEEEDRVEEDGKGRKNEKAGSKRKKSYTSKKSSKQSRKSPGDEDDKCKEENKSSS 180
Db 121 SSEEEDRVEEDGKGRKNEKAGSKRKKSYTSKKSSKQSRKSPGDEDDKCKEENKSSS 180

QY 181 EGGDAGNDTRNTTSDLOKTSSEGT 203
Db 181 EGGDAGNDTRNTTSDLOKTSSEGT 203

RESULT 3
US-09-866-050A-651
; Sequence 651, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 651
; LENGTH: 202
; TYPE: PRT

Query Match 100.0%; Score 1086; DB 12; Length 203;
Best Local Similarity 100.0%; Pred. No. 4.9e-82; Indels 0; Gaps 0;
Matches 203; Conservative 0; Mismatches 0;

QY 1 MARPPREYKAGDLVFAKMKGYPHWPARIDELPEGAVKPPANKYPIFFFGTHETAFGLGPK 60
Db 1 MARPPREYKAGDLVFAKMKGYPHWPARIDELPEGAVKPPANKYPIFFFGTHETAFGLGPK 60

QY 61 DLFPYKEYKDKFGKSNKRGFNEGLWEIENNPVGVKFTGYQAIQQSSSSETEGEGNTADA 120
Db 61 DLFPYKEYKDKFGKSNKRGFNEGLWEIENNPVGVKFTGYQAIQQSSSSETEGEGNTADA 120

QY 121 SSEEEDRVEEDGKGRKNEKAGSKRKKSYTSKKSSKQSRKSPGDEDDKCKEENKSSS 180
Db 121 SSEEEDRVEEDGKGRKNEKAGSKRKKSYTSKKSSKQSRKSPGDEDDKCKEENKSSS 180

QY 181 EGGDAGNDTRNTTSDLOKTSSEGT 203
Db 181 EGGDAGNDTRNTTSDLOKTSSEGT 203

RESULT 4
US-09-925-301-987
; Sequence 987, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 987
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-987

Query Match 82.4%; Score 895; DB 9; Length 172;
Best Local Similarity 98.8%; Pred. No. 2.3e-66;
Matches 169; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 33 PEGAVKPPANKYPIFFFGTHETAFGLGPKDLFPYKEYKDKFGKSNKRGFNEGLWEIENNP 92
Db 2 PRGAVKPSANKYPIFFFGTHETAFGLGPKDLFPYKEYKDKFGKSNKRGFNEGLWEIENNP 61

QY 93 GVKFTGYQAIQQSSSSETEGEGNTADASSSEEGDRVEDGKGRKNEKAGSKRKKSYTS 152
Db 62 GVKFTGYQAIQQSSSSETEGEGNTADASSSEEGDRVEDGKGRKNEKAGSKRKKSYTS 121

QY 153 KSSSKQSRKSPGDEDDKCKEENKSSSEGGDAGNDTRNTTSDLOKTSSEGT 203
Db 122 KSSSKQSRKSPGDEDDKCKEENKSSSEGGDAGNDTRNTTSDLOKTSSEGT 172

RESULT 5
US-09-768-826-60
; Sequence 60, Application US/09768826
; Patent No. US20020012966A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: 18 human secreted proteins
; FILE REFERENCE: PF512P1
; CURRENT APPLICATION NUMBER: US/09/768,826
; CURRENT FILING DATE: 2001-01-25
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2004, 15:28:22 ; Search time 23 Seconds
(without alignments)
648.876 Million cell updates/sec

Title: US-09-787-328B-4
Perfect score: 1086
Sequence: 1 MARPPREYKAGDLVFAKMK.....DAGNDRNTTSDLQKTSECT 203

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 352987 seqs, 73517924 residues

Total number of hits satisfying chosen parameters: 352987

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_AA_New.*
1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pcp.*
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pcp.*
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pcp.*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pcp.*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pcp.*
6: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pcp.*
7: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	486.5	44.8	530	1	PCT-US03-38193-1115
2	486.5	44.8	530	6	US-10-723-860-1115
3	343	31.6	65	6	US-10-425-114A-41803
4	189	17.4	547	6	US-10-745-586-137
5	189	17.4	553	6	US-10-481-779-12
6	186.5	17.2	475	5	US-09-614-150A-2601
7	134.5	12.4	764	7	US-60-490-890-1847
8	134	12.3	765	6	US-10-679-063-25865
9	129.5	11.9	304	6	US-10-739-930-7468
10	128.5	11.8	727	6	US-10-679-063-7384
11	128.5	11.8	764	6	US-10-679-063-20098
12	127.5	11.7	727	6	US-10-679-063-7103
13	127.5	11.7	727	6	US-10-679-063-15050
14	127.5	11.7	764	6	US-10-679-063-7102
15	127.5	11.7	764	6	US-10-679-063-15049
16	127.5	11.7	1360	1	PCT-US03-36702-24
17	127.5	11.7	1360	6	US-10-714-228-24
18	127.5	11.7	1360	7	US-60-490-890-211
19	124.5	11.5	191	6	US-10-679-063-13279
20	122.5	11.3	488	1	PCT-US03-38193-2984
21	122.5	11.3	488	6	US-10-723-860-2984
22	122.5	11.3	488	7	US-60-485-450-1300
23	122.5	11.3	488	7	US-60-495-135-377
24	122.5	11.3	488	7	US-60-495-135-378
25	122.5	11.3	498	7	US-60-531-203-3
26	122.5	11.3	498	7	US-60-531-567-3

27	122.5	11.3	498	7	US-60-531-371-3	Sequence 3, Appli
28	119.5	11.0	231	5	US-09-614-150A-21075	Sequence 21075, A
29	119.5	11.0	315	6	US-10-425-114A-38959	Sequence 38959, A
30	119.5	11.0	315	6	US-10-425-114A-46993	Sequence 46993, A
31	119.5	11.0	315	6	US-10-425-114A-53526	Sequence 53526, A
32	119.5	11.0	315	6	US-10-425-114A-60313	Sequence 60313, A
33	119	11.0	1278	5	US-09-614-150A-42438	Sequence 42438, A
34	118	10.9	312	6	US-10-425-114A-47459	Sequence 47459, A
35	118	10.9	517	6	US-10-425-114A-62674	Sequence 62674, A
36	118	10.9	564	6	US-10-425-114A-42986	Sequence 42986, A
37	117.5	10.8	637	6	US-10-679-063-18043	Sequence 18043, A
38	116.5	10.7	422	6	US-10-739-930-9013	Sequence 9013, Ap
39	116	10.7	281	6	US-10-296-115-918	Sequence 918, App
40	116	10.7	1664	5	US-09-614-150A-2484	Sequence 2484, Ap
41	114	10.5	629	7	US-60-490-890-1888	Sequence 1888, Ap
42	114	10.5	802	7	US-60-490-890-1890	Sequence 1890, Ap
43	114	10.5	1365	7	US-60-490-890-1892	Sequence 1892, Ap
44	113	10.4	658	7	US-60-490-890-2143	Sequence 2143, Ap
45	112	10.3	254	6	US-10-425-114A-61267	Sequence 61267, A

ALIGNMENTS

RESULT 1
PCT-US03-38193-1115
; Sequence 1115, Application PC/TUS0338193
; GENERAL INFORMATION:
; APPLICANT: Ariz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.00PC00
; CURRENT APPLICATION NUMBER: PCT/US03/38193
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1115
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-38193-1115

Query Match	44.8%	Score 486.5;	DB 1;	Length 530;
Best Local Similarity	38.8%	Pred. No. 3.6e-24;		
Matches	109;	Conservative	29;	Mismatches 42; Indels 101; Gaps 7;
Qy	7	REYKAGDLVFAKMGYPHPARIDELPEGAVKPPANKYPIFFPGTHETAPLGPDLPPYK	66	
Db	3	RFKFGDLIFAKMGYPHPARVDEPDGAVAPPTNKLPIFFPGTHETAPLGPDLPPYS	62	
Qy	67	EYKDFGKSNKRKGFNEGLWEIENPNPGVKFTGYQAIQQOS--SSETEGEGGNTA-----	118	
Db	63	ENKEYGKPKRKGNEGLWEIDNNPKVFSQQAATKQSNASSDVEVEKETSVSKEDT	122	
Qy	119	-----DASSEE-----	124	
Db	123	DHEEKASNEEDVTKAVDITTPKAARRGRKRAEKQVETEAGVVVTTATASVNLKVSPPGR	182	
Qy	125	-----EGDRVEEKGKGRK--NEKAGSKR--KKSYSKSSK	157	
Db	183	PAATEVKIPKPRGRPMVKQPCPSSSDIITEEDSKKQGEKQPKKQPKDEEGKEED	242	
Qy	158	QSRKSPGDEDDKCKEENK-----SSSEGGDAGND	188	
Db	243	KPRKEPDKKEGK--KEVESKRNKLAKTGVTSTSDSEEGDD	281	

RESULT 2
US-10-723-860-1115

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2004, 15:26:52 ; Search time 21 Seconds

(without alignments)

929.631 Million cell updates/sec

Title: US-09-787-328B-4

Perfect score: 1086

Sequence: 1 MARPRREYKAGDLVFAKMK.....DAGNDTRNTTSDLQKTSEGT 203

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1086	100.0	203	JC7163	hepatoma-derived g
2	534.5	49.2	240	A55055	hepatoma-derived g
3	532	49.0	237	JC5660	hepatoma-derived g
4	504.5	46.5	669	JC5662	hepatoma-derived g
5	486.5	44.8	530	JC7168	lens epithelium-de
6	337.5	31.1	283	JC5661	hepatoma-derived g
7	193	17.8	1095	J25520	hypothetical prote
8	162.5	15.0	1392	T51947	probable transcrip
9	162	14.9	1445	T50508	hypothetical prote
10	161	14.8	1366	B84924	hypothetical prote
11	155.5	14.3	1261	T48096	hypothetical prote
12	142	13.1	568	T39897	hypothetical prote
13	134.5	12.4	420	T01077	hypothetical prote
14	134.5	12.4	727	S18193	autoantigen NOR-90
15	134.5	12.4	764	S09318	transcription fact
16	134	12.3	765	S22314	transcription fact
17	132	12.2	488	I46014	cyclin II - bovin
18	131	12.1	524	S35551	transcription fact
19	130	12.0	210	T28771	hypothetical prote
20	130	12.0	211	T25911	hypothetical prote
21	128.5	11.8	727	B40439	UBF transcription
22	128.5	11.8	764	A40439	UBF transcription
23	128.5	11.8	1093	T50651	AP-3 complex beta3
24	128.5	11.8	1094	T50651	AP-3 complex beta3
25	127.5	11.7	727	JC5113	ribosomal transcri
26	127.5	11.7	764	JC5112	ribosomal transcri
27	127.5	11.7	1360	JC5939	GRBP-N protein - h
28	127	11.7	532	T06029	hypothetical prote
29	124.5	11.5	1305	T00670	probable inositol

30	122.5	11.3	522	2	C96608	hypothetical prote
31	121.5	11.2	307	2	T04141	histone deacetylase
32	121.5	11.2	902	2	T47966	hypothetical prote
33	120.5	11.1	219	2	T19897	hypothetical prote
34	120.5	11.1	598	2	B40713	cyclin I - human
35	119.5	11.0	2206	2	G71611	hypothetical prote
36	118.5	10.9	751	2	S65469	DNA topoisomerase
37	118.5	10.9	806	2	T23648	hypothetical prote
38	117.5	10.9	1079	2	T38913	translation initia
39	117.5	10.8	327	2	T30072	hypothetical prote
40	117.5	10.8	414	2	A39205	nuclear localizati
41	117.5	10.8	1428	1	ISBYT2	DNA topoisomerase
42	117	10.8	654	2	B71623	knob-associated Hi
43	117	10.8	763	2	T08929	hypothetical prote
44	116	10.7	697	2	T16908	hypothetical prote
45	115.5	10.6	1178	2	S78475	mannosylphosphoryl

ALIGNMENTS

RESULT 1

JC7163

hepatoma-derived growth factor-related protein 3 - human

C;Species: Homo sapiens (man)

C;Date: 04-Mar-2000 #sequence_revision 04-Mar-2000 #text_change 21-Jul-2000

C;Accession: JC7163

R;Ikigame, K.; Yamamoto, M.; Kishima, Y.; Enomoto, H.; Yoshida, K.; Suemura, M.; Kishim

Biochem. Biophys. Res. Commun. 266, 81-87, 1999

A;Title: A new member of a hepatoma-derived growth factor gene family can translocate t

A;Reference number: JC7163; MUID:20050055; PMID:10581169

A;Accession: JC7163

A;Molecule type: mRNA

A;Residues: 1-203 <I>KE>

A;Cross-references: DDBJ:AB029156; NID:96855467; PIDN:BAA90477.1; PID:96855468

C;Comment: This protein, a new member of the hepatoma-derived growth factor gene family

associates to the nucleus and induces cellular growth.

C;Genetics:

A;Gene: hrp-3

A;Map position: 15q25

C;Keywords: differentiation; glycoprotein; growth factor; nucleus

Query Match 100.0%; Score 1086; DB 2; Length 203;

Best Local Similarity 100.0%; Pred. No. 2,7e-69;

Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARPRREYKAGDLVFAKMKGYPHWPARIDELPEGAVKPPANKYPIFFGTHETAFILGPK 60

Db 1 MARPRREYKAGDLVFAKMKGYPHWPARIDELPEGAVKPPANKYPIFFGTHETAFILGPK 60

Qy 61 DLPPYKEYDKPKSKNRKGFNEGLWEIENNPVGKVTGYQAIQQSSSETEGEGGNTADA 120

Db 61 DLPPYKEYDKPKSKNRKGFNEGLWEIENNPVGKVTGYQAIQQSSSETEGEGGNTADA 120

Qy 121 SSEEEDRVVEEDGKGRKNEKAGSKRKYTSKKSKQSKRSPGDDDDKCKEENKSSS 180

Db 121 SSEEEDRVVEEDGKGRKNEKAGSKRKYTSKKSKQSKRSPGDDDDKCKEENKSSS 180

Qy 181 EGGDAGNDTRNTTSDLQKTSEGT 203

Db 181 EGGDAGNDTRNTTSDLQKTSEGT 203

RESULT 2

A55055

hepatoma-derived growth factor - human

C;Species: Homo sapiens (man)

C;Date: 18-Nov-1994 #sequence_revision 18-Nov-1994 #text_change 20-Jun-2000

C;Accession: A55055

R;Nakamura, H.; Izumoto, Y.; Kambe, H.; Kuroda, T.; Mori, T.; Kawamura, K.; Yamamoto, H

J. Biol. Chem. 269, 25143-25149, 1994

A;Title: Molecular cloning of complementary DNA for a novel human hepatoma-derived grow

A;Reference number: A55055; MUID:95014294; PMID:7929202

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2004, 15:26:51 ; Search time 18 Seconds
(without alignments)
530.357 Million cell updates/sec

Title: US-09-787-328B-4

Perfect score: 1086

Sequence: 1 MARPREYKAGDLVFAKMK.....DAGNDTRNTSDLQKTSEGT 203

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	534.5	49.2	240	1 HDGF_HUMAN	P51858 homo sapien
2	532	49.0	237	1 HDGF_MOUSE	P51859 mus musculus
3	134.5	12.4	764	1 UBFI_HUMAN	P17480 homo sapien
4	134	12.3	765	1 UBFI_MOUSE	P25976 mus musculus
5	132	12.2	488	1 T2FA_BOVIN	Q28092 bos taurus
6	131	12.1	524	1 T2FA_XENLA	Q04870 xenopus lae
7	128.5	11.8	764	1 UBFI_RAT	P25977 rattus norv
8	128.5	11.8	1094	1 A3BI_HUMAN	O00203 homo sapien
9	127.5	11.7	406	1 NSBI_MOUSE	O91335 mus musculus
10	127.5	11.7	1360	1 MSH6_HUMAN	P52701 homo sapien
11	126.5	11.6	875	1 ARS2_MOUSE	O99m76 mus musculus
12	124	11.4	1358	1 MSH6_MOUSE	P54276 mus musculus
13	122.5	11.3	488	1 HD42_HUMAN	Q92769 homo sapien
14	120.5	11.1	598	1 CYLI_HUMAN	P35663 homo sapien
15	118.5	10.9	1079	1 IFP2_SCHPO	Q10251 saccharomyc
16	117.5	10.8	414	1 NSRI_YEAST	P27476 saccharomyc
17	117.5	10.8	1394	1 CNG4_BOVIN	Q28181 bos taurus
18	117.5	10.8	1428	1 TOP2_YEAST	P06786 saccharomyc
19	116.5	10.7	488	1 HD42_CHICK	P56519 gallus gall
20	116	10.7	556	1 HIR3_HUMAN	Q9bw71 homo sapien
21	116	10.7	697	1 SSRP_CABEL	P41848 caenorhabdi
22	115.5	10.6	1178	1 MN4_YEAST	P36044 saccharomyc
23	113.5	10.5	416	1 CRIC_BERST	Q92ppl berberis st
24	113.5	10.5	1082	1 A3BI_HUMAN	Q13367 homo sapien
25	113	10.4	500	1 GAR2_SCHPO	P41891 schizosacch
26	113	10.4	676	1 S521_HUMAN	Q96mu7 homo sapien
27	112.5	10.4	797	1 VG48_HSVSA	Q01033 herpesvirus
28	112.5	10.4	908	1 SRCA_RABIT	P13666 oryctolagus
29	111.5	10.3	480	1 HD12_XENLA	Q42227 xenopus lae
30	111.5	10.3	482	1 HD42_HUMAN	Q13547 homo sapien
31	111.5	10.3	488	1 HD42_MOUSE	P70288 mus musculus
32	111.5	10.3	633	1 MLH_TETTH	P40631 tetrahymena
33	111.5	10.3	1105	1 A3BI_MOUSE	Q921t1 mus musculus

RESULT 1

ID	HDGF_HUMAN	STANDARD;	PRT;	240 AA.
AC	P51858;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Hepatoma-derived growth factor (HDGF) (High-mobility group protein 1-like 2) (HMG-1L2).			
GN	HDGF OR HMG1L2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 4-24.			
RC	TISSUE=Hepatoma;			
RX	MEDLINE=95014294; PubMed=7929202;			
RA	Nakamura H., Izumoto Y., Kambe H., Kuroda T., Mori T., Kawamura K., Yamamoto H., Kishimoto T.;			
RT	"Molecular cloning of complementary DNA for a novel human hepatoma-derived growth factor. Its homology with high mobility group-1 protein."			
RL	J. Biol. Chem. 269:25143-25149(1994).			

ALIGNMENTS

34	111	10.2	282	1 NSBI_HUMAN	P82970 homo sapien
35	111	10.2	480	1 HD41_CHICK	P56517 gallus gall
36	111	10.2	634	1 KNOB_PLAFA	P09346 plasmodium
37	111	10.2	876	1 ARS2_HUMAN	Q9bxs5 homo sapien
38	110.5	10.2	480	1 HD11_XENLA	Q91895 xenopus lae
39	110	10.1	687	1 DSPP_RAT	Q62598 rattus norv
40	109.5	10.1	415	1 CRIC_RICCO	P93508 ricinus com
41	109.5	10.1	483	1 CBFS_YEAST	P33322 saccharomyc
42	109	10.0	348	1 CYL2_HUMAN	Q14093 homo sapien
43	109	10.0	914	1 PBPA_BACSU	P39793 bacillus su
44	109	10.0	934	1 DSPP_MOUSE	P97399 mus musculus
45	108.5	10.0	416	1 CRIC_NICPL	Q40401 nicotiana p

CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- TISSUE SPECIFICITY: Ubiquitous.
CC -|- SIMILARITY: Contains 1 FWWP domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; D16431; BAA03903.1; -
 DR EMBL; BC018991; AAH18991.1; -
 DR PIR; A55055; A55055
 DR Genew; HGNC:4856; HDGF.
 DR MIM; 300043; -
 DR GO; GO:0005737; C:cytoplasm; TAS.
 DR GO; GO:0005615; C:extracellular space; TAS.
 DR GO; GO:0008201; F:heparin binding activity; TAS.
 DR GO; GO:0008283; P:cell proliferation; TAS.
 DR InterPro; IPR000313; PWMF_domain.
 DR Pfam; PF00855; PWMF; 1.
 DR SMART; SM00293; PWMF; 1.
 DR PROSITE; PS00812; PWMF; 1.
 KW Growth factor; Heparin-binding.
 FT DOMAIN 12 69 PWMF.
 FT DOMAIN 213 223 POLY-GLU.
 SQ SEQUENCE 240 AA; 26788 MW; DD60D9203BDD4B34 CRC64;

Query Match 49.2%; Score 534.5; DB 1; Length 240;
 Best Local Similarity 51.4%; Pred. No. 1.7e-30;
 Matches 114; Conservative 25; Mismatches 58; Indels 25; Gaps 6;
 QY 5 RPEYKAGDLVPAKMGYPHPARIDELPEGAVKPPANKYPIFFGTHTETAPLGPOLFP 64
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 6 ROKEYKCGDLVPAKMGYPHPARIDENPEAAVKSTANKYQVFRTGTHTETAPLGPOLFP 65
 QY 65 YKEYKDKGKSKNRKGFNEGLWEIENPVKFTGYQATQQSSSE-----TPEG-- 114
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 66 YEESKEKFGPKNRKGFSEGLWEIENPVTKASGYQSSQSCVPEPEPEPAEGDGDK 125
 QY 115 -QNTADASSEEGDVEEDGKRGKNEKAGSRKKS---YTSKSKSKSRKSPGDEDDKD 170
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 126 KGN-REGSSDEEG-KLVIDEPAKNEKAGLRAGDILLESFKPKPAENPEGEKEAA 183
 QY 171 CHE-----EENKSSSEGGDAGNDTRNTTSDLOKTSGET 203
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 184 TLEVERPLPMEVKNSTSEPGSGRGPPQEEEEEDEEEBEAT 225

RESULT 2

HDGF MOUSE
 ID HDGF MOUSE STANDARD; PRT; 237 AA.
 AC P51859; Q9CYA4; Q9JK87;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hepatoma-derived growth factor (HDGF).
 GN HDGF OR TDRM1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Testis;
 RX MEDLINE=97445118; PubMed=9299445;
 RA Izumoto Y., Kuroda T., Harada H., Kishimoto T., Nakamura H.;
 RT "Hepatoma-derived growth factor belongs to a gene family in mice
 RT showing significant homology in the amino terminus."
 RL Biochem. Biophys. Res. Commun. 238:26-32(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Thymus;
 RA Zhao Y., Chen W., Wang Y.;
 RT "Cloning of novel gene related to thymus development."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Mateau Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayaishizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver, and Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenfer C.F., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Skalski D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: Heparin-binding protein, with mitogenic activity for
 CC fibroblasts (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN TESTIS AND SKELETAL
 CC MUSCLE, TO INTERMEDIATE EXTENTS IN HEART, BRAIN, LUNG, LIVER, AND
 CC KIDNEY, AND TO A MINIMAL EXTENT IN SPLEEN.
 CC -!- SIMILARITY: Contains 1 PWMF domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; D63707; BAA09838.1; -
 DR EMBL; AF251787; AAF65469.1; -
 DR EMBL; AK017863; BAB30979.1; -
 DR EMBL; BC005713; AAH05713.1; -
 DR EMBL; BC021654; AAH21654.1; -
 DR PIR; JC5660; JC5660.
 DR MGD; MGI:1194494; Hdgf.
 DR InterPro; IPR000313; PWMF_domain.
 DR Pfam; PF00855; PWMF; 1.
 DR SMART; SM00293; PWMF; 1.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2004, 15:26:51 ; Search time 35 Seconds
(without alignments)

1496.705 Million cell updates/sec

Title: US-09-787-328B-4

Perfect score: 1086

Sequence: 1 MARPPREYKAGDLVFAKMK.....DAGNDTRNTSDLQKTSEGT 203

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL 23.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1086	100.0	203	4 Q9Y3E1	Q9Y3E1 homo sapien
2	1059.5	97.6	202	11 Q9JMG7	Q9JMG7 mus musculus
3	1053.5	97.0	202	11 Q923W4	Q923W4 rattus norv
4	1050.5	96.7	202	11 Q9D2M7	Q9D2M7 mus musculus
5	1045.5	96.3	205	11 Q8BQ69	Q8BQ69 mus musculus
6	851.5	78.4	220	11 Q8BR62	Q8BR62 mus musculus
7	532	49.0	237	11 Q8BPG7	Q8BPG7 mus musculus
8	530.5	48.8	239	6 Q9XSK7	Q9XSK7 bos taurus
9	530	48.8	237	11 Q8VHK7	Q8VHK7 rattus norv
10	530	48.8	237	11 Q923W3	Q923W3 rattus norv
11	524	48.3	670	4 Q9BW08	Q9BW08 homo sapien
12	512.5	47.2	669	11 Q925G1	Q925G1 rattus norv
13	504.5	46.5	669	11 Q35540	Q35540 mus musculus
14	504.5	46.5	678	11 Q99192	Q99192 mus musculus
15	486.5	44.8	333	4 Q9N2I3	Q9N2I3 homo sapien
16	486.5	44.8	530	4 Q75475	Q75475 homo sapien

17	485.5	44.7	333	4 Q95368	Q95368 homo sapien
18	485.5	44.7	530	4 Q9UER6	Q9UER6 homo sapien
19	485	44.7	331	11 Q99JF7	Q99JF7 mus musculus
20	485	44.7	528	11 Q95JF8	Q95JF8 mus musculus
21	481	44.3	283	13 Q8AVR7	Q8AVR7 xenopus lae
22	467.5	43.0	146	11 Q9CT03	Q9CT03 mus musculus
23	411	37.8	235	6 Q9XSK6	Q9XSK6 bos taurus
24	361.5	33.3	251	4 Q9NUE8	Q9NUE8 homo sapien
25	354.5	32.6	251	4 Q96MJ6	Q96MJ6 homo sapien
26	342.5	31.5	302	4 Q8NPR8	Q8NPR8 homo sapien
27	342	31.5	296	11 Q8VIH2	Q8VIH2 rattus norv
28	341	31.4	283	11 Q35539	Q35539 mus musculus
29	269.5	24.8	483	6 Q8MJG1	Q8MJG1 bos taurus
30	251.5	23.2	258	11 Q9ESF6	Q9ESF6 mus musculus
31	238	21.9	50	4 Q8N4N4	Q8N4N4 homo sapien
32	193	17.8	832	5 Q95Q90	Q95Q90 caenorhabdi
33	193	17.8	835	5 Q95Q91	Q95Q91 caenorhabdi
34	189	17.4	547	4 Q9BXX2	Q9BXX2 homo sapien
35	188.5	17.4	546	11 Q922P9	Q922P9 mus musculus
36	186.5	17.2	475	5 Q9VAA9	Q9VAA9 drosophila
37	179.5	16.5	546	11 Q9CYQ1	Q9CYQ1 mus musculus
38	167	15.4	523	4 Q9BT11	Q9BT11 homo sapien
39	162.5	15.0	1392	10 Q9XER9	Q9XER9 arabidopsis
40	162	14.9	1445	10 Q9LEY4	Q9LEY4 arabidopsis
41	161	14.8	1366	10 Q8GYT7	Q8GYT7 arabidopsis
42	160.5	14.8	1449	10 Q8H3E3	Q8H3E3 oryza sativ
43	160	14.7	33	11 Q8C9E2	Q8C9E2 mus musculus
44	155.5	14.3	1261	10 Q9LYB5	Q9LYB5 arabidopsis
45	142	13.1	568	3 Q94312	Q94312 schizosacch

ALIGNMENTS

RESULT 1

ID	Q9Y3E1	PRELIMINARY;	PRT;	203 AA.
AC	Q9Y3E1;			
DT	01-NOV-1999	(TrEMBLrel. 12, Created)		
DT	01-NOV-1999	(TrEMBLrel. 12, Last sequence update)		
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)		
DE	CGI-142 protein (Hypothetical protein FLJ10418) (HRP-3) (Hepatoma-derived growth factor 2).			
GN	HRP-3 OR HDGF2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Lin W.-C.;			
RT	"Comparative gene cloning: Identification of novel human genes with Caenorhabditis elegans proteome as template.";			
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y., Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H., Sugawara M., Watanabe S., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S., Yamamoto J., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K., Ninomiya K., Iwayanagi T.;			
RT	"NEDO human cDNA sequencing project.";			
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=20050055; PubMed=10581169;			
RX	Ikegami K., Yamamoto M., Kishima Y., Enomoto H., Yoshida K., Suemura M., Kishimoto T., Nakamura H.;			
RA	"A new member of a hepatoma-derived growth factor gene family can translocate to the nucleus.";			
RL	Biochem. Biophys. Res. Commun. 266:81-87(1999).			


```
[4]
RN  SEQUENCE FROM N.A.
RC  TISSUE=uterus;
RA  Strausberg R.;
RL  Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN  [5]
RA  Yu L., Fu Q., Tu Q.;
RT  Cloning of a novel human cDNA which is a homolog to mouse hepatoma-
RT  derived growth factor (mHGF) and termed hHGF2.;
RL  Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF151900; AAD34137.1; -.
DR  EMBL; AK001280; BAA91597.1; -.
DR  EMBL; AB029156; BAA90477.1; -.
DR  EMBL; BC015483; AAH15483.1; -.
DR  EMBL; AF110642; AA027003.1; -.
DR  InterPro; IPR000313; PWMP_domain.
DR  Pfam; PF00855; PWMP; 1.
DR  PROSITE; PS50812; PWMP; 1.
KW  Hypothetical protein.
SQ  SEQUENCE 203 AA; 22620 MW; 0B660D665F01659C CRC64;

Query Match      100.0%; Score 1086; DB 4; Length 203;
Best Local Similarity 100.0%; Pred. No. 6.7e-76;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  1 MARPPREYKAGDLVFAKMGYPHPWPARIDELPEGAVKPPANKYPPIFFFGTHETAFLGPK 60
DB  1 MARPPREYKAGDLVFAKMGYPHPWPARIDELPEGAVKPPANKYPPIFFFGTHETAFLGPK 60

QY  61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGKFTGYQAIQQOSSSETEGEGNTADA 120
DB  61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGKFTGYQAIQQOSSSETEGEGNTADA 120

QY  121 SSEEGRDVEEDGKGRKNEKAGSKRKYTSKSSKQSRKSPGDEDDKCKEENKSSS 180
DB  121 SSEEGRDVEEDGKGRKNEKAGSKRKYTSKSSKQSRKSPGDEDDKCKEENKSSS 180

QY  181 EGGDAGNDRNTTSDLOKTSEGT 203
DB  181 EGGDAGNDRNTTSDLOKTSEGT 203

RESULT 2
Q9JMG7 PRELIMINARY; PRT; 202 AA.
ID  Q9JMG7
AC  Q9JMG7;
DT  01-OCT-2000 (TrEMBLrel. 15, Created)
DT  01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE  Hepatoma-derived growth factor-related protein HRP-3.
GN  HDGFRP3 OR HRP-3.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]_TaxID=10090;
RP  SEQUENCE FROM N.A.
RC  TISSUE=Testis;
RA  Ikegama K., Nakamura H.;
RL  Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Testis;
RX  MEDLINE=20050055; PubMed=10581169;
RA  Ikegama K., Yamamoto M., Kishina Y., Enomoto H., Yoshida K.,
RA  Suetomura M., Kishimoto T., Nakamura H.;
RT  "A new member of a hepatoma-derived growth factor gene family can
RT  translocate to the nucleus.";
RL  Biochem. Biophys. Res. Commun. 266:81-87 (1999).
DR  EMBL; AB029493; BAA90478.1; -.
DR  MGI; MGI:1352760; Hdgfrp3.
DR  InterPro; IPR000313; PWMP_domain.

Query Match      97.0%; Score 1053.5; DB 11; Length 202;
Best Local Similarity 97.5%; Pred. No. 2.1e-73;
Matches 198; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY  1 MARPPREYKAGDLVFAKMGYPHPWPARIDELPEGAVKPPANKYPPIFFFGTHETAFLGPK 60
DB  1 MARPPREYKAGDLVFAKMGYPHPWPARIDELPEGAVKPPANKYPPIFFFGTHETAFLGPK 60

QY  61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGKFTGYQAIQQOSSSETEGEGNTADA 120
DB  61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGKFTGYQAIQQOSSSETEGEGNTADA 120

QY  121 SSEEGRDVEEDGKGRKNEKAGSKRKYTSKSSKQSRKSPGDEDDKCKEENKSSS 180
DB  121 SSEEGRDVEEDGKGRKNEKAGSKRKYTSKSSKQSRKSPGDEDDKCKEENKSSS 180

QY  181 EGGDAGNDRNTTSDLOKTSEGT 203
DB  181 EGGDAGNDRNTTSDLOKTSEGT 203

RESULT 4
Q9D2M7
```

```

ID Q9D2M7 PRELIMINARY; PRT; 202 AA.
AC Q9D2M7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hepatoma-derived growth factor, related protein 3.
GN HDGFRP3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Iehli Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schrim L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barch G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Satō K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Iyoy-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR ENBL; AK019487; BAB31754.1; -.
DR DB; AK044858; BAC3123.1; -.
DR MCD; MGI:1352760; Hdgfrp3.
DR IncerPro; IPR000313; PWWP_domain.
DR Pfam; PF00855; PWWP; 1.
DR SMART; SM00293; PWWP; 1.
DR PROSITE; PS0812; PWWP; 1.
SQ SEQUENCE 202 AA; 22430 MW; FCDB31BD2BEA8B6 CRC64;

Query Match 96.7%; Score 1050.5; DB 11; Length 202;
Best Local Similarity 97.0%; Pred. No. 3.5e-73;
Matches 197; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

Qy 1 MARPRREYKAGDLVFAKMGYPHPWPARIDELPEGAVKPPANKYPIFFGTHETAFILGPK 60
Db 1 MARPRREYKAGDLVFAKMGYPHPWPARIDELPEGAVKPPANKYPIFFGTHETAFILGPK 60

Qy 61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGKFTGYQAIQQSSSETEGEGNTADA 120
Db 61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGKFTGYQAIQQSSSETEGEGNTADA 120

Qy 121 SSEEGRVDEEDGKGRKNEKAGSKKSYTSKKSKQSRKSPGDEDDKCKEENKSS 180
Db 121 SSEEGRV-EDGKGRKNEKAGSKKSYTSKKSKQSRKSPGDEDDKCKEENKSS 179

Qy 181 EGGDAGNDTRNTTSDLQKTS 203
Db 180 EGGDAGNDTRNTTADLQKAG 202

Query Match 78.4%; Score 851.5; DB 11; Length 220;
Best Local Similarity 90.5%; Pred. No. 7.1e-58;
Matches 162; Conservative 4; Mismatches 12; Indels 1; Gaps 1;

```

```

RESULT 5
Q8BQ69 PRELIMINARY; PRT; 205 AA.
AC Q8BQ69;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hepatoma-derived growth factor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Dorsal root ganglion;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR ENBL; AK051401; BAC34628.1; -.
DR EMBL; AK051401; BAC34628.1; -.
DR SRA; E051401.1; BAC34628.1; -.
SQ SEQUENCE 205 AA; 22819 MW; E2ABAE3FBB31BD2B CRC64;

Query Match 96.3%; Score 1045.5; DB 11; Length 205;
Best Local Similarity 97.0%; Pred. No. 8.7e-73;
Matches 196; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

Qy 1 MARPRREYKAGDLVFAKMGYPHPWPARIDELPEGAVKPPANKYPIFFGTHETAFILGPK 60
Db 1 MARPRREYKAGDLVFAKMGYPHPWPARIDELPEGAVKPPANKYPIFFGTHETAFILGPK 60

Qy 61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGKFTGYQAIQQSSSETEGEGNTADA 120
Db 61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGKFTGYQAIQQSSSETEGEGNTADA 120

Qy 121 SSEEGRVDEEDGKGRKNEKAGSKKSYTSKKSKQSRKSPGDEDDKCKEENKSS 180
Db 121 SSEEGRV-EDGKGRKNEKAGSKKSYTSKKSKQSRKSPGDEDDKCKEENKSS 179

Qy 181 EGGDAGNDTRNTTSDLQKTS 202
Db 180 EGGDAGNDTRNTTADLQKAG 201

Query Match 96.3%; Score 1045.5; DB 11; Length 205;
Best Local Similarity 97.0%; Pred. No. 8.7e-73;
Matches 196; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

RESULT 6
Q8BR62 PRELIMINARY; PRT; 220 AA.
AC Q8BR62;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hepatoma-derived growth factor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR ENBL; AK045542; BAC32413.1; -.
DR SRA; E045542.1; BAC32413.1; -.
SQ SEQUENCE 220 AA; 25094 MW; 96C692B40976C79F CRC64;

Query Match 78.4%; Score 851.5; DB 11; Length 220;
Best Local Similarity 90.5%; Pred. No. 7.1e-58;
Matches 162; Conservative 4; Mismatches 12; Indels 1; Gaps 1;

```

Qy	1	MARPRPREYKAGDLVFAKMGYPHPWPARI	DELPEGAVKPPANKYPIFFPGTHETAFLGPK	60
Dy	1	MARPRPREYKAGDLVFAKMGYPHPWPARI	DELPEGAVKPPANKYPIFFPGTHETAFLGPK	60
Qy	61	DLPPYKEYKDFGKGNKRGKGFNEGLWEI	NNPGVKFTGYQAIQQOSSSETEGEGNTADA	120
Dy	61	DLPPYKEYKDFGKGNKRGKGFNEGLWEI	NNPGVKFTGYQTIQQOSSSETEGEGNTADA	120
Qy	121	SSEEGDRVVEEDGKGRKNEKAGSKRK	SYTSKSSKQSRKSPGDEDDKCKEENKSS	179
Dy	121	SSEEGDRV-EDGKGRKNEKGSKRK	SYTSKSSKQSRKSQETKMIKAKRRRTKAA	178
RESULT 7				
Q8BPG7	ID	Q8BPG7	PRELIMINARY;	PRT; 237 AA.
AC	Q8BPG7			
DT	01-MAR-2003	(T=EMBLrel. 23, Created)		
DT	01-MAR-2003	(T=EMBLrel. 23, Last sequence update)		
DT	01-MAR-2003	(T=EMBLrel. 23, Last annotation update)		
DE	Hepatoma-derived growth factor.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Body;			
RX	MEDLINE=22354683; PubMed=12466851;			
RA	The FANTOM Consortium,			
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;			
RT	"Analysis of the mouse transcriptome based on functional annotation of			
RT	60,770 full-length cDNAs."			
RL	Nature 420:563-573 (2002).			
DR	EMBL; AK076021; BAC36126.1; --			
SQ	SEQUENCE 237 AA; 26270 MW; BBA0CF574DA4733F CRC64;			
Query Match 49.0%; Score 532; DB 11; Length 237;				
Best Local Similarity 53.9%; Pred. No. 2.5e-33;				
Matches 111; Conservative 29; Mismatches 38; Indels 28; Gaps 6				
Qy	5	RPRYKAGDLVFAKMGYPHPWPARI	DELPEGAVKPPANKYPIFFPGTHETAFLGPKDLFP	64
Dy	6	RQKEYKCGDLVFAKMGYPHPWPARI	DEMPAAVKSTANKYQVFFPGTHETAFLGPKDLFP	65
Qy	65	YKEYKDFGKGNKRGKGFNEGLWEI	NNPGVKFTGYQAIQQOSSSETEGEG 114	
Dy	66	YEEKKFGKGNKRGKGFNEGLWEI	NNPTVKAGSYQSSQKSCAAAEPEVEPEAHEGDGDK 125	
Qy	115	GNTADASSSEEGDRVVEEDGKGRKNEKAGSKRK	SYTSKSSKQSRKSPGDEDDKCKE- 173	
Dy	126	KGSAEGSSDEG-KLVIDPAKEKNEKGLTKRRAGDVLDS	PPRKES-GDHEED-KEI 182	
Qy	174	-----EENKSSSEGGDAG 186		
Dy	183	AALGERPLPVEVEKNSTPSE-PDSG 207		
RESULT 8				
Q9XSK7	ID	Q9XSK7	PRELIMINARY;	PRT; 239 AA.
AC	Q9XSK7			
DT	01-NOV-1999	(T=EMBLrel. 12, Created)		
DT	01-NOV-1999	(T=EMBLrel. 12, Last sequence update)		
DT	01-OCT-2002	(T=EMBLrel. 22, Last annotation update)		
DE	Hepatoma derived growth factor.			
GN	HDGF.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI TaxID=9913;			